

# Random coding exponents galore via decoupling

Naresh Sharma

Tata Institute of Fundamental Research

Mumbai 400005, India

Email: `nsharma@tifr.res.in`

September 30, 2015

## Abstract

A missing piece in quantum information theory, with very few exceptions, has been to provide the random coding exponents for quantum information-processing protocols. We remedy the situation by providing these exponents for a variety of protocols including those at the top of the family tree of protocols. Our line of attack is to provide an exponential bound on the decoupling error for a restricted class of completely positive maps where a key term in the exponent is in terms of a Rényi  $\alpha$ -information-theoretic quantity for any  $\alpha \in (1, 2]$ . Among the protocols covered are fully quantum Slepian-Wolf, quantum state merging, quantum state redistribution, quantum/classical communication across channels with side information at the transmitter with or without entanglement assistance, and quantum communication across broadcast channels.

## 1 Introduction

Analysis of optimal resources needed/generated in an information-processing protocol is one of the holy grails of information theory [1, 2, 3, 4, 5]. Nice answers in terms of information-theoretic quantities are obtained, in general, for large copies such as of inputs and channel uses. One part in establishing these answers is the achievability that says that for resources arbitrarily close to the optimal, there exists a protocol accomplishing the task with arbitrarily small error.

Achievability proofs come in various flavors and we list some of them but not in the chronological order. One way is via the law of large numbers (or typicality) that involves making statements for large copies. Another way is via the smooth information-theoretic quantities that are defined in terms of a semi-definite program (see Refs. [6, 7] and references therein). This method has the advantage that one can make statements for any number of copies and it matches the optimal answer for large number of copies using

the law of large numbers. A third way has been via the random coding exponents, i.e., one makes statements for any number of copies by obtaining an exponential bound on the error of the protocol. In many comparisons with the second method, this method provides stronger bounds and was pioneered by Gallager who obtained such bounds for the classical capacity [8, 9]. Yet another method has been via the optimal terms in the asymptotic expansions of the rate at which the resources are generated or used and this was pioneered by Strassen [10].

It is the Gallager's approach that would be further investigated in this paper. If one scours the literature on the random coding exponents for quantum protocols, one finds that not much work has been done on this topic. Indeed, apart from Burnashev and Holevo [11], Holevo [12], and Hayashi [13], no other work, to the best of the author's knowledge, provides random coding exponents for the quantum protocols. (Exponential bounds on the error for the Schumacher compression can be obtained without much difficulty leveraging the analysis for the classical source compression [3].) Burnashev and Holevo [11] provide the reliability function (loosely defined as the best exponent one could get for large number of copies [9]) for sending classical information across the quantum channel for the case of pure states, and Holevo [12] extends it for the case of commuting density matrices. Hayashi provides a random coding exponent for the same protocol for general density matrices but his exponent when specialized to classical does not match with Gallager's [12, 8].

Quantum information theory is much richer than the classical and with a plethora of protocols (one can just glance at the family tree of quantum protocols [14, 15] to appreciate this), it is not just important to provide the random coding exponents but, if possible, also a unified approach to get these exponents for a variety of protocols.

Where would such a unified approach come from? An answer lies in decoupling, a phenomenon where random evolution of a part of the quantum system would, on the average, make it decouple from the other part. That decoupling would be useful for quantum error correction was first observed by Schumacher and Nielsen [16]. It has subsequently been recognized as a building block in quantum information theory (see Refs. [17, 18] and references therein).

The decoupling theorem quantifies the average error between the state, part of which is randomly evolved, and the completely decoupled state, and is now known in various versions. We go through some of them not necessarily in the chronological order. The one provided by Hayden *et al* [19] gives a bound in terms of dimensions of the quantum systems involved and this, with an appeal to typicality for large copies, yields the optimal answers – similar approach is followed by Abeyesinghe *et al* [20]. Dupuis *et al* provide another version that gives a bound in terms of smooth entropies [21].

Another version by Dupuis gives an exponential bound for any number of copies and the exponent has two Rényi 2-conditional entropies: first one is computed using the density matrix that is evolved and the second one is computed using the Choi-Jamiołkowski representation of a map [17].

Since this version gives an exponential bound, it seems close to the stated purpose of

this paper but it is not quite there simply because for the random coding exponents, we shall need the first term to be in terms of Rényi  $\alpha$ -conditional entropies for  $\alpha$  arbitrarily close to 1. It is not necessary to strengthen the second term that determines the rate.

Could there be a way of modifying Dupuis' bound? This paper stems from asking this question, answers it in the affirmative, and then applies the new version to obtain the random coding exponents for a variety of protocols. In particular, we are able to replace the first term by a Rényi  $\alpha$ -conditional entropy for all  $\alpha \in (1, 2]$  (although adding some inconsequential terms in the process). We do this by leveraging ideas from the independent works of Dupuis and Hayashi [17, 3].

Some of the protocols we analyze are at the top of the family tree of protocols and the author didn't encounter any protocol that could be analyzed by other versions of the decoupling theorem but not from the version provided in this paper. For the protocols analyzed, the application of our version of the decoupling theorem is, in some cases, but not always, inspired by the application of other versions of the decoupling theorem.

We don't address how close the exponent in the proposed bounds might be to the reliability function. There is, however, one resemblance between the exponents we obtain and the reliability function for the classical case (in certain regimes), which is that in both the cases, it is in terms of Rényi  $\alpha$ -information-theoretic quantities.

The structure of the paper is as follows. Section 2 provides the notation and definitions used throughout this paper. Section 3 provides a new version of the decoupling theorem. (There is a more general version provided as well in Appendix C although we don't use it!) The subsequent sections apply this version to various protocols. Following protocols are analyzed: Schumacher compression, fully quantum Slepian-Wolf, fully quantum reverse Shannon, quantum state merging, quantum/classical communication across channels with side information at the transmitter with or without entanglement assistance, entanglement-assisted classical communication, quantum state redistribution, quantum communication across broadcast channels, and destroying correlations by adding classical randomness. The lemmas are provided in the appendix so as to not interrupt the flow.

## 2 Notation and Preliminaries

Let  $\mathcal{H}_A$  be the Hilbert space associated with the quantum system  $A$ . We shall confine ourselves to the finite dimensional Hilbert spaces in this paper and  $|A|$  denotes the dimension of  $\mathcal{H}_A$ .  $A \cong B$  implies that  $|A| = |B|$ . For a system  $A$ , we denote  $A^n$  to be a quantum system described by  $\bigotimes_{i=1}^n \mathcal{H}_{A_i}$ , where  $A_i \cong A$ ,  $i = 1, \dots, n$ . Let  $L(\mathcal{H}_A, \mathcal{H}_B)$  be the set of all matrices from  $\mathcal{H}_A$  to  $\mathcal{H}_B$  and  $L(\mathcal{H}_A)$  denotes  $L(\mathcal{H}_A, \mathcal{H}_A)$ . Let  $\text{Herm}(\mathcal{H}_A)$ ,  $\text{Pos}(\mathcal{H}_A) \subseteq L(\mathcal{H}_A)$  be the set of Hermitian and positive semidefinite matrices respectively. Let  $D(\mathcal{H}_A) \subseteq \text{Pos}(\mathcal{H}_A)$  be the set of unit trace matrices and  $D_{\leq}(\mathcal{H}_A) \subseteq \text{Pos}(\mathcal{H}_A)$  be the set of matrices with trace not greater than 1. Let  $\nu_{\sigma^A}$  be the number of distinct eigenvalues of  $\sigma^A \in \text{Herm}(\mathcal{H}_A)$ . For  $\rho^A, \sigma^A \in \text{Herm}(\mathcal{H}_A)$ , let  $\{\rho^A \geq \sigma^A\}$  denote the projector onto the subspace spanned by the eigenvectors corresponding to the non-negative eigenval-

ues of  $\rho^A - \sigma^A$ . Let  $X \cdot \rho \equiv X\rho X^\dagger$ . For  $X \in L(\mathcal{H}_A, \mathcal{H}_B)$  (also denoted as  $X^{A \rightarrow B}$ ), the trace norm,  $\|X\|_1$ , is the sum of its singular values. The Fidelity between  $\rho, \sigma \in \text{Pos}(\mathcal{H}_A)$  is  $F(\rho, \sigma) \equiv \|\sqrt{\rho}\sqrt{\sigma}\|_1$ .

Let  $\mathbb{U}(A)$  be a Unitary 2-design on a quantum system  $A$  (see Ref. [17] and references therein). For a function  $f : \mathbb{U}(A) \rightarrow L(\mathcal{H}_E)$ ,  $E_U f(U)$  denotes the expectation taken over a random Unitary  $U$  distributed uniformly on  $\mathbb{U}(A)$ .

Let  $|\Phi\rangle^{AA'}$  be the maximally entangled state (MES) on  $AA'$ , i.e., for  $A \cong A'$ , orthonormal bases  $\{|i\rangle^A\}$  and  $\{|i\rangle^{A'}\}$ ,  $|\Phi\rangle^{AA'} \equiv |A|^{-1/2} \sum_{i=1}^{|A|} |i\rangle^A |i\rangle^{A'}$ . Let the maximally mixed state in  $\mathcal{H}_A$  be denoted by  $\pi^A \equiv \mathbb{1}^A/|A|$ , where  $\mathbb{1}^A$  is the Identity matrix. The zero matrix (with all entries as zero) is denoted by  $0$ .

A matrix  $V^{A \rightarrow B}$  is an isometry if either  $V^\dagger V = \mathbb{1}$  or  $VV^\dagger = \mathbb{1}$ , and is a partial isometry if its singular values are either 0 or 1. A full-rank partial isometry  $V^{A \rightarrow B}$  has rank  $\min\{|A|, |B|\}$ .

The Kronecker delta function is  $\delta_{j,k} = 1$  if  $j = k$ , and 0 otherwise. The indicator function  $\text{ind}_{\text{condition}} = 1$  if condition is true, and 0 otherwise. The partial trace over  $B$  of  $\rho^{AB} \in L(\mathcal{H}_A \otimes \mathcal{H}_B)$  is denoted by either  $\text{Tr}_B \rho^{AB}$  or  $\rho^A$ . For a pure state  $|\Psi\rangle^{AB}$ ,  $\Psi^{AB} = |\Psi\rangle\langle\Psi|^{AB}$ , and it does not necessarily imply that  $\Psi^A$  is also a pure state. All the logarithms in this paper are to the base 2 and  $\exp(x)$  denotes  $2^x$ ,  $x \in \mathbb{R}$ . We define  $\Xi(\varepsilon) \equiv \sqrt{\varepsilon(2 + \varepsilon + 2\sqrt{1 + \varepsilon})}$  for  $\varepsilon \geq 0$ .

With an abuse of notation, we call a weighted sum of exponentially decaying terms also as exponential decay, i.e., for  $x, \alpha_i, \beta_i > 0, i = 1, \dots, n, n$  finite, we call  $\sum_{i=1}^n \beta_i \exp\{-\alpha_i x\}$  as exponentially decaying with  $x$ . All the error bounds that we provide in this paper can be put in this form.

## 2.1 Super-operators

A super-operator  $\mathcal{T}^{A \rightarrow B}$  is a map from  $L(\mathcal{H}_A) \rightarrow L(\mathcal{H}_B)$ . Important classes include completely positive maps  $\mathcal{T}^{A \rightarrow B}$ , which map  $\text{Pos}(\mathcal{H}_A \otimes \mathcal{H}_R)$  to  $\text{Pos}(\mathcal{H}_B \otimes \mathcal{H}_R)$  for any ancilla  $R$ , and completely positive and trace preserving (cptp) maps which are completely positive and have an additional property that the trace is preserved.

The Choi-Jamiołkowski representation of a map  $\mathcal{T}^{A \rightarrow E}$  is given by  $\omega_{\mathcal{T}}^{EA'} \equiv \mathcal{T}^{A \rightarrow E}(\Phi^{AA'})$ . To a completely positive map, we associate a quantity  $\Theta(\mathcal{T})$  defined as the negative of the Rényi old 2-conditional entropy (defined in Section 2.2) and is given by

$$\Theta(\mathcal{T}) \equiv -H_2^{\text{old}}(A'|E)_{\omega_{\mathcal{T}}^{EA'}}. \quad (1)$$

Concatenation of two maps, i.e.,  $\mathcal{E}$  followed by  $\mathcal{D}$  is denoted by  $\mathcal{D} \circ \mathcal{E}$ , and with a slight abuse of notation, for an isometry  $V$  and a map  $\mathcal{E}$ ,  $\mathcal{E} \circ V(\rho)$  denotes  $\mathcal{E}(V \cdot \rho)$ , and  $V \circ \mathcal{E}(\rho)$  denotes  $V \cdot \mathcal{E}(\rho)$ .

We now define three maps. For  $\sigma^{AB} \in L(\mathcal{H}_A \otimes \mathcal{H}_B)$ ,  $\mathcal{Q}_A(\sigma^{AB}) \equiv |A| \text{Tr}_A \sigma^{AB} (\sigma^{AB})^\dagger - \sigma^B (\sigma^B)^\dagger$ . For  $\rho, \sigma \in \text{Pos}(\mathcal{H}_A)$ , the spectral decomposition  $\sigma = \sum_{i=1}^{\nu_\sigma} \lambda_i \Pi_i$ , where  $\lambda_i$ 's are all distinct and  $\Pi_i$ 's are projectors, a pinching map in the eigenbasis of  $\sigma$  is defined as

$\mathcal{M}_\sigma(\rho) \equiv \sum_{i=1}^{\nu_\sigma} \Pi_i \rho \Pi_i$ . Let  $W^{A \rightarrow B}$ ,  $|B| \leq |A|$ , be a full-rank partial isometry. Then a compressive map  $\mathcal{C}_W^{A \rightarrow B}$  is defined as  $\mathcal{C}_W(\rho^A) \equiv W \rho^A W^\dagger + [\text{Tr}(\mathbb{1}^A - W^\dagger W) \rho^A] \pi^B$ .

**Definition 1** (Class-1 maps). A map  $\mathcal{T}^{A \rightarrow E}$  is said to be in class-1 if it is completely positive and for any  $\sigma \in \mathcal{L}(\mathcal{H}_A)$ ,  $\mathbb{E}_U \|\mathcal{T}(U \cdot \sigma)\|_1 \leq \|\sigma\|_1$ .

Note that all cptp maps fall under class-1. Another set of completely positive maps under class-1 are those with  $\text{Tr} \mathcal{T}(\mathbb{1}^A) = |A|$  (see Lemma 25 for proof). An example of such a map (taken from Ref. [17]) that we shall use later in the paper is given by

$$\mathcal{T}_W^{A \rightarrow B}(\sigma^A) \equiv \frac{|A|}{|B|} (W^{A \rightarrow B} \cdot \sigma^A), \quad (2)$$

where  $W^{A \rightarrow B}$ ,  $|A| \geq |B|$ , is a full-rank partial isometry.

## 2.2 Information-theoretic quantities

The quantum relative entropy from  $\rho$  to  $\sigma$  is given by  $D(\rho \parallel \sigma) \equiv \text{Tr} \rho (\log \rho - \log \sigma)$ , the von Neumann entropy of  $\rho^A \in \mathcal{D}(\mathcal{H}_A)$  is given by  $H(A)_\rho \equiv -\text{Tr} \rho^A \log \rho^A$ . For a tripartite state  $\rho^{ABC}$ , the conditional entropy of  $A$  given  $B$  is given by  $H(A|B)_\rho \equiv H(AB)_\rho - H(B)_\rho$ , the conditional mutual information between  $A$  and  $B$  given  $C$  is  $I(A : B|C)_\rho \equiv H(A|C)_\rho - H(A|BC)_\rho$ , and the coherent information is given by  $I(A \rangle B)_\rho \equiv -H(A|B)_\rho$ . The Rényi generalizations of the quantum relative entropy can be done in various ways and we mention two prominent candidates.

**Definition 2** (Rényi entropies). For  $\alpha \in (0, 2] \setminus \{1\}$ , from  $\rho$  to  $\sigma$ , the quasi old  $\alpha$ -relative entropy is given by  $Q_\alpha^{\text{old}}(\rho \parallel \sigma) \equiv \text{Tr} \rho^\alpha \sigma^{1-\alpha}$ , and the quasi sandwiched  $\alpha$ -relative entropy (proposed independently in Refs. [22, 23]) is given by  $Q_\alpha^{\text{sand}}(\rho \parallel \sigma) \equiv \text{Tr} \left( \sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}} \right)^\alpha$ . The Rényi old (sandwiched)  $\alpha$ -relative entropy from  $\rho$  to  $\sigma$  is given by

$$D_\alpha^{\text{old (sand)}}(\rho \parallel \sigma) \equiv \frac{1}{\alpha - 1} \log Q_\alpha^{\text{old (sand)}}(\rho \parallel \sigma), \quad \alpha \in (0, 2] \setminus \{1\}. \quad (3)$$

We can extend these definitions to include  $\alpha = 1$  by taking limits and we drop the subscript and the superscript. The Rényi  $\alpha$ -conditional entropies of  $A$  given  $B$  are defined as

$$H_\alpha^{\text{type}}(A|B)_\rho \equiv - \inf_{\sigma^B \in \mathcal{D}(\mathcal{H}_B)} D_\alpha^{\text{type}}(\rho^{AB} \parallel \mathbb{1}^A \otimes \sigma^B) \quad (4)$$

$$\downarrow H_\alpha^{\text{type}}(A|B)_\rho \equiv -D_\alpha^{\text{type}}(\rho^{AB} \parallel \mathbb{1}^A \otimes \rho^B), \quad (5)$$

where ‘type’ is ‘old’ or ‘sand’.

It follows from Refs. [24, 25, 22, 23] that for  $\alpha \in (0, 2] \setminus \{1\}$  and a cptp map  $\mathcal{E}$ ,  $D_\alpha^{\text{type}}(\rho \parallel \sigma) \geq D_\alpha^{\text{type}}[\mathcal{E}(\rho) \parallel \mathcal{E}(\sigma)]$ .

There are duality relations known for a tripartite pure state  $\Psi^{ABC}$ . One such example is  $H_\alpha^{\text{sand}}(A|B)_\Psi + H_\alpha^{\text{sand}}(A|C)_\Psi = 0$ ,  $\tilde{\alpha} = 1/\alpha$ ,  $\alpha \in [0.5, 1) \cup (1, 2]$ . See Ref. [26] and

references therein for a complete list of duality relations. In the remainder of the paper, the ‘type’ superscript is dropped, and it implies that the expression holds for either one and one could pick one’s favorite. For example,  $D_\alpha(\rho\|\sigma)$  denotes either  $D_\alpha^{\text{old}}(\rho\|\sigma)$  or  $D_\alpha^{\text{sand}}(\rho\|\sigma)$ . Furthermore, while invoking the above duality relations, since there are many options, we also drop the downarrow superscript from the conditional entropies and assume that appropriate superscript is implicitly assumed and  $\tilde{\alpha}$  is assumed to be an appropriate function of  $\alpha$  depending on the type of conditional entropies involved.

### 3 Yet another version of the decoupling theorem with a useful Rényiification

In this section, we provide a version of the decoupling theorem where the crucial term in the exponent is in terms of a Rényi  $\alpha$ -information-theoretic quantity for  $\alpha \in (1, 2]$  instead of just  $\alpha = 2$  as provided in Ref. [17].

We leverage ideas from the independent works of Dupuis and Hayashi and in particular Theorem 3.7 in Ref. [17] and Lemma 9.2 in Ref. [3].

**Theorem 1.** *Let  $\rho^{AR} \in \mathcal{D}(\mathcal{H}_{AR})$  and  $\mathcal{T}^{A \rightarrow E}$  be a class-1 map. Then for  $\alpha \in (1, 2]$ , a random Unitary  $U$  acting on  $A$ , we have for any  $\sigma^R \in \mathcal{D}(\mathcal{H}_R)$ ,*

$$\begin{aligned} \mathbb{E}_U \left\| \mathcal{T}(U \cdot \rho^{AR}) - \omega_{\mathcal{T}}^E \otimes \rho^R \right\|_1 \\ \leq 4 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ \log \nu_{\sigma^R} + D_\alpha(\rho^{AR} \| \mathbb{1}^A \otimes \sigma^R) + \Theta(\mathcal{T}) \right] \right\}. \end{aligned} \quad (6)$$

*In particular, for  $n$  copies, a random Unitary  $U$  acting on  $A^n$ , and a class-1 map  $\mathcal{T}^{A^n \rightarrow E}$ , we have*

$$\begin{aligned} \mathbb{E}_U \left\| \mathcal{T} [U \cdot (\rho^{AR})^{\otimes n}] - \omega_{\mathcal{T}}^E \otimes (\rho^R)^{\otimes n} \right\|_1 \\ \leq 4 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |R| \log(n + 1) - n H_\alpha(A|R)_\rho + \Theta(\mathcal{T}) \right] \right\}. \end{aligned} \quad (7)$$

*Proof.* For a  $\zeta > 0$ , let  $\Pi^{AR} \equiv \{\mathcal{M}_{\mathbb{1}^A \otimes \sigma^R}(\rho^{AR}) \geq \zeta \mathbb{1}^A \otimes \sigma^R\}$ ,  $\hat{\Pi}^{AR} \equiv \mathbb{1}^{AR} - \Pi^{AR}$ ,  $\mu_1 \equiv \omega_{\mathcal{T}}^E \otimes \text{Tr}_A \{\Pi^{AR} \rho^{AR}\}$ , and  $\mu_2 \equiv \omega_{\mathcal{T}}^E \otimes \text{Tr}_A \{\hat{\Pi}^{AR} \rho^{AR}\}$ . Note that  $\mu_1 + \mu_2 = \omega_{\mathcal{T}}^E \otimes \rho^R$ . We now have

$$\begin{aligned} \mathbb{E}_U \left\| \mathcal{T}(U \cdot \rho^{AR}) - \omega_{\mathcal{T}}^E \otimes \rho^R \right\|_1 &= \mathbb{E}_U \left\| \mathcal{T} [U \cdot (\Pi^{AR} \rho^{AR})] - \mu_1 + \mathcal{T} [U \cdot (\hat{\Pi}^{AR} \rho^{AR})] - \mu_2 \right\|_1 \\ &\leq \mathbb{E}_U \left\| \mathcal{T} [U \cdot (\Pi^{AR} \rho^{AR})] - \mu_1 \right\|_1 + \mathbb{E}_U \left\| \mathcal{T} [U \cdot (\hat{\Pi}^{AR} \rho^{AR})] - \mu_2 \right\|_1, \end{aligned} \quad (8)$$

where we have used the triangle inequality.



We attack the first term.

$$\mathbb{E}_U \|\mathcal{T}[U \cdot (\Pi^{AR} \rho^{AR})] - \mu_1\|_1 \leq \mathbb{E}_U \|\mathcal{T}[U \cdot (\Pi^{AR} \rho^{AR})]\|_1 + \|\mu_1\|_1 \quad (9)$$

$$\leq 2 \mathbb{E}_U \|\mathcal{T}[U \cdot (\Pi^{AR} \rho^{AR})]\|_1 \quad (10)$$

$$\leq 2 \|\Pi^{AR} \rho^{AR}\|_1 \quad (11)$$

$$\leq 2\zeta^{\frac{1-\alpha}{2}} \exp \left\{ \frac{\alpha-1}{2} D_\alpha(\rho^{AR} \| \mathbb{1}^A \otimes \sigma^R) \right\}, \quad (12)$$

where the first inequality follows from the triangle inequality, the second inequality follows from the convexity of the trace norm to have

$$\|\mu_1\|_1 = \|\mathbb{E}_U \{\mathcal{T}[U \cdot (\Pi^{AR} \rho^{AR})]\}\|_1 \leq \mathbb{E}_U \|\mathcal{T}[U \cdot (\Pi^{AR} \rho^{AR})]\|_1, \quad (13)$$

the third inequality follows from the definition of class-1 maps, the fourth inequality follows from Lemma 28 (proved by Hayashi [3]).

We now attack the second term. Let  $\Delta_U \equiv \mathcal{T}[U \cdot (\hat{\Pi}^{AR} \rho^{AR})] - \mu_2$ , and  $\theta^E \in \mathcal{D}(\mathcal{H}_E)$  be such that  $\Theta(\mathcal{T}) = -D_2^{\text{old}}(\omega_{\mathcal{T}}^{EA'} \| \theta^E \otimes \mathbb{1}^{A'})$ . We now have

$$\mathbb{E}_U \|\Delta_U^{ER}\|_1 \leq \mathbb{E}_U \sqrt{\text{Tr}[(\theta^E)^{-1} \otimes (\sigma^R)^{-1}] \Delta_U \Delta_U^\dagger} \quad (14)$$

$$\leq \sqrt{\text{Tr}[(\theta^E)^{-1} \otimes (\sigma^R)^{-1}] \mathbb{E}_U \{\Delta_U \Delta_U^\dagger\}} \quad (15)$$

$$\leq \sqrt{\frac{|A|^2 \text{Tr} \left\{ (\theta^E)^{-1} \text{Tr}_{A'} (\omega_{\mathcal{T}}^{EA'})^2 \right\} \text{Tr} \left\{ (\sigma^R)^{-1} \text{Tr}_A \hat{\Pi}^{AR} (\rho^{AR})^2 \hat{\Pi}^{AR} \right\}}{|A|^2 - 1}} \quad (16)$$

$$\leq \sqrt{\frac{\nu_{\sigma^R} \zeta |A|^2 \exp \{\Theta(\mathcal{T})\}}{|A|^2 - 1}}, \quad (17)$$

where the first inequality follows since for any matrix  $\Upsilon$  and a density matrix  $\kappa$  (with appropriate dimensions),  $\|\Upsilon\|_1 \leq \sqrt{\text{Tr} \kappa^{-1} \Upsilon \Upsilon^\dagger}$ , the second inequality follows from the concavity of  $x \rightarrow \sqrt{x}$ , the third inequality follows from Lemma 27, and the last inequality follows from Lemma 29 (proved by Hayashi [3]). We now have

$$\begin{aligned} \mathbb{E}_U \|\mathcal{T}(U \cdot \rho^{AR}) - \omega_{\mathcal{T}}^E \otimes \rho^R\|_1 \\ \leq 2\zeta^{\frac{1-\alpha}{2}} \exp \left\{ \frac{\alpha-1}{2} D_\alpha(\rho^{AR} \| \mathbb{1}^A \otimes \sigma^R) \right\} + \sqrt{\frac{\nu_{\sigma^R} \zeta |A|^2 \exp \{\Theta(\mathcal{T})\}}{|A|^2 - 1}}. \end{aligned} \quad (18)$$

Note that  $\zeta$  is a free parameter and a convenient upper bound for

$$\min_{\zeta} \left( x \zeta^{\frac{1-\alpha}{2}} + y \zeta^{1/2} \right) \quad (19)$$

is obtained by choosing  $\zeta = (xy^{-1})^{\frac{2}{\alpha}}$ . Making that choice by feeding in appropriate values of  $x, y$ , taking  $|A|^2/(|A|^2 - 1) \leq 4/3$ , noting that for  $\alpha \in (1, 2]$ ,  $2^{1/\alpha}(4/3)^{(\alpha-1)/(2\alpha)} \leq 2$ , we get

$$\begin{aligned} \mathbb{E}_U \left\| \mathcal{T}(U \cdot \rho^{AR}) - \omega_{\mathcal{T}}^E \otimes \rho^R \right\|_1 \\ \leq 4 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ \log \nu_{\sigma^R} + D_{\alpha}(\rho^{AR} \| \mathbb{1}^A \otimes \sigma^R) + \Theta(\mathcal{T}) \right] \right\}. \end{aligned} \quad (20)$$

Note that this is a convenient upper bound and while one could further optimize the choice of  $\zeta$ , for  $\alpha$  near 1, the above bound is near the optimal.

For  $n$  copies, a random Unitary over  $A^n$ , and a class-1 map  $\mathcal{T}^{A^n \rightarrow E}$ , using (20), we have

$$\begin{aligned} \mathbb{E}_U \left\| \mathcal{T} [U \cdot (\rho^{AR})^{\otimes n}] - \omega_{\mathcal{T}}^E \otimes (\rho^R)^{\otimes n} \right\|_1 \\ \leq 4 \min_{\sigma^R} \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ \nu_{(\sigma^R)^{\otimes n}} + D_{\alpha} [(\rho^{AR})^{\otimes n} \| \mathbb{1}^{A^n} \otimes (\sigma^R)^{\otimes n}] + \Theta(\mathcal{T}) \right] \right\} \\ \leq 4 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |R| \log(n+1) - nH_{\alpha}(A|R)_{\rho} + \Theta(\mathcal{T}) \right] \right\}, \end{aligned} \quad (21)$$

where the first inequality follows from (20) and making a (possibly suboptimal) choice of a product state, and the second inequality follows since we have used a convenient upper bound that for any  $\sigma^R \in \mathcal{D}(\mathcal{H}_R)$ ,  $\log \nu_{(\sigma^R)^{\otimes n}} \leq |R| \log(n+1)$  (see Theorem 11.1.1 in Ref. [1] or Lemma 3.7 in Ref. [3]) and we choose  $\sigma^R$  to be the one that minimizes  $D_{\alpha}(\rho^{AR} \| \mathbb{1}^A \otimes \sigma^R)$ . QED.  $\square$

We now have the following corollary of Theorem 1.

**Corollary 2.** For  $i = 1, \dots, K$ , let  $\mathcal{T}_i^{A^n \rightarrow E_i}$  be class-1 maps, and  $\rho_i^{AR_i} \in \mathcal{D}(\mathcal{H}_{AR_i})$ . Then there exists a Unitary  $U$  over  $A^n$  such that for all  $i = 1, \dots, K$ , and  $n \in \mathbb{N}$ ,

$$\begin{aligned} \left\| \mathcal{T}_i [U \cdot (\rho_i^{AR_i})^{\otimes n}] - \omega_{\mathcal{T}_i}^{E_i} \otimes (\rho_i^{R_i})^{\otimes n} \right\|_1 \\ \leq 4K \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |R_i| \log(n+1) - nH_{\alpha}(A|R_i)_{\rho_i} + \Theta(\mathcal{T}_i) \right] \right\}. \end{aligned} \quad (22)$$

*Proof.* It follows from Theorem 1 that for all  $i = 1, \dots, K$ ,

$$\begin{aligned} \mathbb{E}_U \left\| \mathcal{T}_i [U \cdot (\rho_i^{AR_i})^{\otimes n}] - \omega_{\mathcal{T}_i}^{E_i} \otimes (\rho_i^{R_i})^{\otimes n} \right\|_1 \\ \leq 4 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |R_i| \log(n+1) - nH_{\alpha}(A|R_i)_{\rho_i} + \Theta(\mathcal{T}_i) \right] \right\}. \end{aligned} \quad (23)$$

We now invoke Lemma I.7 in Ref. [17] to arrive at the claim. (Note that Lemma I.7 in Ref. [17] stipulates a multiplying factor of  $K + 1$  instead of  $K$  but it can be easily strengthened.)  $\square$

It is possible to provide a unified theorem that yields both Theorem 1 and Lemma 9.2 in Ref. [3] as special cases. We do that in Theorem 33 (see Appendix C) and we note that although we provide this unified theorem, we don't use it for the protocols treated later in the paper!



## 4 Schumacher compression

**Definition 3.** A  $(\rho, \text{error}, n)$  **Schumacher compression** protocol consists of  $n$  copies of  $\rho^A$  (with a purification  $\Psi^{AR}$ ), Alice applying an encoding cptp map  $\mathcal{E} : A^n \rightarrow B$ , and Bob applying a decoding cptp map  $\mathcal{D} : B \rightarrow \tilde{A}^n$  such that for  $\rho^{\tilde{A}^n R^n} \equiv \mathcal{D}^{B \rightarrow \tilde{A}^n} \circ \mathcal{E}^{A^n \rightarrow B} [(\Psi^{AR})^{\otimes n}]$ ,

$$\left\| \rho^{\tilde{A}^n R^n} - (\Psi^{\tilde{A}R})^{\otimes n} \right\|_1 \leq \text{error}. \quad (24)$$

$(\log |B|)/n$  is called the **compression rate** of the protocol. A real number  $\mathcal{R}_C$  is called an **achievable rate** if there exist, for  $n \rightarrow \infty$ , Schumacher compression protocols with compression rate approaching  $\mathcal{R}_C$  and the error approaching 0.

**Theorem 3** (Schumacher, 1995 [27]). *The smallest achievable rate for Schumacher compression is given by  $H(A)_\rho$ .*

We prove the following theorem.

**Theorem 4.** *For any  $n \in \mathbb{N}$ , there exists a  $(\rho, \text{error}, n)$  Schumacher compression protocol such that for any  $\delta > 0$ ,*

$$\frac{\log |B|}{n} = |R| \frac{\log(n+1)}{n} + H_{\tilde{\alpha}}(A)_\Psi + \delta. \quad (25)$$

and the error approaches 0 exponentially in  $n$ .

*Proof.* Consider a full-rank partial isometry  $W^{A^n \rightarrow B}$ ,  $|B| \leq |A|^n$ . Then, using Theorem 1, there exists a Unitary  $U$  over  $A^n$ ,

$$\begin{aligned} & \left\| \text{Tr}_B \circ \mathcal{T}_W^{A^n \rightarrow B} [U \cdot (\Psi^{AR})^{\otimes n}] - (\Psi^R)^{\otimes n} \right\|_1 \\ & \leq 4 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |R| \log(n+1) - nH_\alpha(A|R)_\Psi + \Theta(\text{Tr}_B \circ \mathcal{T}_W) \right] \right\} \\ & = 4 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |R| \log(n+1) + nH_{\tilde{\alpha}}(A)_\Psi - \log |B| \right] \right\} \equiv \varepsilon_n, \end{aligned} \quad (26)$$

where we have used  $\Theta(\text{Tr}_B \circ \mathcal{T}_W) = -\log |B|$  from Lemma 21. We claim using Lemma 31 that there exists a Unitary  $V^{A^n \rightarrow A^n}$  such that

$$\left\| W^\dagger \cdot \mathcal{T}_W^{A^n \rightarrow B} [U \cdot (\Psi^{AR})^{\otimes n}] - V \cdot (\Psi^{AR})^{\otimes n} \right\|_1 \leq \Xi(\varepsilon_n), \quad (27)$$

and hence, using monotonicity of the trace norm under cptp maps ( $\mathcal{C}_W$  in this case),

$$\left\| \mathcal{T}_W^{A^n \rightarrow B} [U \cdot (\Psi^{AR})^{\otimes n}] - \mathcal{C}_W [V \cdot (\Psi^{AR})^{\otimes n}] \right\|_1 \leq \Xi(\varepsilon_n), \quad (28)$$

or

$$\left\| W^\dagger \cdot \mathcal{C}_W [V \cdot (\Psi^{AR})^{\otimes n}] - W^\dagger \cdot \mathcal{T}_W^{A^n \rightarrow B} [U \cdot (\Psi^{AR})^{\otimes n}] \right\|_1 \leq \Xi(\varepsilon_n). \quad (29)$$

Define a partial isometry  $W_2^{A^n \rightarrow B}$  as  $W_2 \equiv WV$  and note that  $\mathcal{C}_{W_2}(\sigma^{AR}) = \mathcal{C}_W(V \cdot \sigma^{AR})$ . We now have

$$\begin{aligned} & \left\| W_2^\dagger \cdot \mathcal{C}_{W_2} [(\Psi^{AR})^{\otimes n}] - (\Psi^{AR})^{\otimes n} \right\|_1 \\ &= \left\| V^\dagger W^\dagger \cdot \mathcal{C}_W [V \cdot (\Psi^{AR})^{\otimes n}] - (\Psi^{AR})^{\otimes n} \right\|_1 \end{aligned} \quad (30)$$

$$= \left\| W^\dagger \cdot \mathcal{C}_W [V \cdot (\Psi^{AR})^{\otimes n}] - V \cdot (\Psi^{AR})^{\otimes n} \right\|_1 \quad (31)$$

$$\begin{aligned} &\leq \left\| W^\dagger \cdot \mathcal{C}_W [V \cdot (\Psi^{AR})^{\otimes n}] - W^\dagger \cdot \mathcal{T}^{A^n \rightarrow B} [U \cdot (\Psi^{AR})^{\otimes n}] \right\|_1 + \\ &\quad \left\| W^\dagger \cdot \mathcal{T}^{A^n \rightarrow B} [U \cdot (\Psi^{AR})^{\otimes n}] - V \cdot (\Psi^{AR})^{\otimes n} \right\|_1 \end{aligned} \quad (32)$$

$$\leq 2\Xi(\varepsilon_n), \quad (33)$$

where we have used the triangle inequality, (27), and (29). It is now clear that Alice applies  $\mathcal{C}_{W_2}$  and Bob applies the isometry  $W_2^\dagger$ . The claim now follows readily.  $\square$

**Remark:** This is not the best exponent for this protocol and one can get the exponent that matches with the classical case (see Prob. 2.15 in Ref. [28]) when specialized and this can be construed from the treatment in Ref. [3]. Our purpose of stating the above proof is that the ideas would prove useful for other protocols later in this paper since it is based on decoupling.

## 5 Fully quantum Slepian-Wolf (FQSW)

**Definition 4.** A  $(\Psi, \text{error}, n)$  FQSW protocol consists of  $n$  copies of a pure state  $|\Psi\rangle^{ABR}$  shared between with Alice ( $A$ ) and Bob ( $B$ ), and reference system ( $R$ ), Alice applying an encoding cftp map  $\mathcal{E} : A^n \rightarrow A_1 A_2$ , quantum communication across a noiseless quantum channel from Alice to Bob  $\mathcal{I}^{A_2 \rightarrow B_2}$ , and Bob applying a decoding cftp map  $\mathcal{D} : B_2 B^n \rightarrow B_1 \tilde{B}_3^n B_3^n$  such that for

$$\rho^{A_1 B_1 \tilde{B}_3^n B_3^n R^n} \equiv \mathcal{D}^{B_2 B^n \rightarrow B_1 \tilde{B}_3^n B_3^n} \circ \mathcal{I}^{A_2 \rightarrow B_2} \circ \mathcal{E}^{A^n \rightarrow A_1 A_2} [(\Psi^{ABR})^{\otimes n}], \quad (34)$$

$$\left\| \rho^{A_1 B_1 \tilde{B}_3^n B_3^n R^n} - \Phi^{A_1 B_1} \otimes (\Psi^{\tilde{B}_3 B_3 R})^{\otimes n} \right\|_1 \leq \text{error}. \quad (35)$$

The number  $(\log |A_2|)/n$  is called the **quantum communication rate** and  $(\log |A_1|)/n$  is called the **entanglement gain rate** of the protocol.

A pair of real numbers  $(\mathcal{R}_Q, \mathcal{R}_E)$  is called an **achievable rate pair** if there exist, for  $n \rightarrow \infty$ , FQSW protocols with quantum communication rate approaching  $\mathcal{R}_Q$ , entanglement gain rate approaching  $\mathcal{R}_E$ , and error approaching 0.

The achievable rates are described by the following theorem.

**Theorem 5** (Abeyesinghe et al, 2009 [20]). The following rates are achievable for the FQSW:

$$\mathcal{R}_Q > \frac{1}{2} I(A : R)_\Psi \quad \text{and} \quad \mathcal{R}_E < \mathcal{R}_Q + H(A|R)_\Psi. \quad (36)$$

Our goal in the remainder of this section is to provide the achievability of the above rate region with error decaying to 0 exponentially in  $n$ .

**Theorem 6.** *For any  $n \in \mathbb{N}$ , there exists a  $(\Psi, \text{error}, n)$  FQSW protocol for any  $\alpha \in (1, 2]$ , and  $\delta_1, \delta_2 > 0$ , such that*

$$\frac{\log |A_2|}{n} = \frac{1}{2} \left[ H_{\tilde{\alpha}}(A)_{\Psi} - H_{\alpha}(A|R)_{\Psi} \right] + (|B| + 1)|R| \frac{\log(n+1)}{2n} + \frac{\delta_1 + \delta_2}{2}, \quad (37)$$

$$\frac{\log |A_1|}{n} = \frac{\log |A_2|}{n} + H_{\alpha}(A|R)_{\Psi} - |R| \frac{\log(n+1)}{n} - \delta_2, \quad (38)$$

and the error approaches 0 exponentially in  $n$ .

*Proof.* Let  $W : A^n \rightarrow A_1 A_2$  be a full-rank partial isometry with  $|A_1| |A_2| \leq |A|$ . Then, using Corollary 2, we claim that there exists a Unitary  $U$  over  $A^n$  such that for  $\alpha \in (1, 2]$ ,

$$\begin{aligned} & \left\| \text{Tr}_{A_1 A_2} \circ \mathcal{T}_W^{A^n \rightarrow A_1 A_2} [U \cdot (\Psi^{ABR})^{\otimes n}] - (\Psi^{BR})^{\otimes n} \right\|_1 \\ & \leq 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |B| |R| \log(n+1) - n H_{\alpha}(A|BR)_{\Psi} + \Theta(\text{Tr}_{A_1 A_2} \circ \mathcal{T}_W) \right] \right\} \\ & = 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |B| |R| \log(n+1) + n H_{\tilde{\alpha}}(A)_{\Psi} - \log |A_1| |A_2| \right] \right\} \equiv \varepsilon_n, \end{aligned} \quad (39)$$

and

$$\begin{aligned} & \left\| \text{Tr}_{A_2} \circ \mathcal{T}_W^{A^n \rightarrow A_1 A_2} [U \cdot (\Psi^{AR})^{\otimes n}] - \pi^{A_1} \otimes (\Psi^R)^{\otimes n} \right\|_1 \\ & \leq 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |R| \log(n+1) - n H_{\alpha}(A|R)_{\Psi} + \log \frac{|A_1|}{|A_2|} \right] \right\} \equiv \vartheta_n. \end{aligned} \quad (40)$$

It follows from (40) and Lemma 31 that there exists an isometry  $\tilde{U}^{A_2 B^n \rightarrow B_1 \tilde{B}_3^n B_3^n}$  such that

$$\left\| \tilde{U} \cdot \left\{ \mathcal{T}_W^{A^n \rightarrow A_1 A_2} [U \cdot (\Psi^{ABR})^{\otimes n}] \right\} - \Phi^{A_1 B_1} \otimes (\Psi^{\tilde{B}_3^n B_3^n R})^{\otimes n} \right\|_1 \leq \Xi(\vartheta_n). \quad (41)$$

It follows from (39) and Lemma 31 that there exists a Unitary  $V^{A^n \rightarrow A^n}$  such that

$$\begin{aligned} \Xi(\varepsilon_n) & \geq \left\| W^{\dagger} \cdot \mathcal{T}_W^{A^n \rightarrow A_1 A_2} [U \cdot (\Psi^{ABR})^{\otimes n}] - V \cdot (\Psi^{ABR})^{\otimes n} \right\|_1 \\ & \geq \left\| \mathcal{T}_W^{A^n \rightarrow A_1 A_2} [U \cdot (\Psi^{ABR})^{\otimes n}] - \mathcal{C}_W [V \cdot (\Psi^{ABR})^{\otimes n}] \right\|_1 \\ & = \left\| \tilde{U} \cdot \left\{ \mathcal{T}_W^{A^n \rightarrow A_1 A_2} [U \cdot (\Psi^{ABR})^{\otimes n}] \right\} - \tilde{U} \cdot \left\{ \mathcal{C}_W [V \cdot (\Psi^{ABR})^{\otimes n}] \right\} \right\|_1, \end{aligned} \quad (42)$$

where the second inequality follows using the monotonicity and noting that

$$\mathcal{C}_W \{ W^{\dagger} \cdot \mathcal{T}_W^{A^n \rightarrow A_1 A_2} [U \cdot (\Psi^{ABR})^{\otimes n}] \} = \mathcal{T}_W^{A^n \rightarrow A_1 A_2} [U \cdot (\Psi^{ABR})^{\otimes n}], \quad (43)$$

and the last equality is true since  $(\tilde{U})^\dagger \tilde{U} = \mathbb{1}^{A_2 B^n}$ . Lastly, we use the triangle inequality, (41), and (42) to claim that

$$\left\| \tilde{U}^{B_2 B^n \rightarrow B_1 \tilde{B}_3^n B_3^n} \cdot \{ \mathcal{I}^{A_2 \rightarrow B_2} \circ \mathcal{C}_W [V \cdot (\Psi^{ABR})^{\otimes n}] \} - \Phi^{A_1 B_1} \otimes (\Psi^{\tilde{B}_3^n B_3^n R})^{\otimes n} \right\|_1 \leq \Xi(\varepsilon_n) + \Xi(\vartheta_n). \quad (44)$$

It follows that the protocol consists of Alice applying  $\mathcal{C}_W^{A \rightarrow A_1 A_2} \circ V^{A^n}$ , and Bob applies  $\tilde{U}$ , albeit on  $B_2 B^n$  instead of  $A_2 B^n$ .

It is now clear that if, for  $\alpha \in (1, 2]$ ,  $\delta_1, \delta_2 > 0$ ,

$$\frac{\log |A_1|}{n} + \frac{\log |A_2|}{n} = H_{\tilde{\alpha}}(A)_\Psi + |B||R| \frac{\log(n+1)}{n} + \delta_1, \quad (45)$$

$$\frac{\log |A_1|}{n} - \frac{\log |A_2|}{n} = H_\alpha(A|R)_\Psi - |R| \frac{\log(n+1)}{n} - \delta_2, \quad (46)$$

then the error decays exponentially in  $n$  to zero.

The claim of the theorem now follows and we exhaust the entire achievable rate region as stipulated by Theorem 5.

Note that in view of the trivial protocol where one qubit transmitted across a noiseless qubit channel from Alice and Bob generates one EPR pair shared by Alice and Bob (the reverse implication is not true), it makes sense to keep the quantum communication as small as possible, which is accomplished by making  $\alpha$  close to 1, and  $\delta_1, \delta_2$  close to 0.  $\square$

## 6 Fully quantum reverse Shannon (FQRS)

The following definition is from Ref. [20].

**Definition 5.** A  $(\Psi, \text{error}, n)$  FQRS protocol consists of  $n$  copies of a pure state  $|\Psi\rangle^{AA'}$  (both  $A$  and  $A'$  held by Alice), a MES  $\Phi^{A_1 B_1}$  shared between Alice ( $A_1$ ) and Bob ( $B_1$ ), a cptp map  $\mathcal{N}^{A' \rightarrow B}$  with Stinespring representation  $V_N^{A' \rightarrow BE}$  and  $|\Psi\rangle^{ABE} = V_N^{A' \rightarrow BE} |\Psi\rangle^{AA'}$ , Alice applying an encoding cptp map  $\mathcal{E} : A'^n A_1 \rightarrow A_2 E^n$ , quantum communication across a noiseless quantum channel from Alice to Bob  $\mathcal{I}^{A_2 \rightarrow B_2}$ , and Bob applying a decoding cptp map  $\mathcal{D} : B_1 B_2 \rightarrow B^n$  such that for

$$\rho^{A^n B^n E^n} \equiv \mathcal{D}^{B_1 B_2 \rightarrow B^n} \circ \mathcal{I}^{A_2 \rightarrow B_2} \circ \mathcal{E}^{A'^n A_1 \rightarrow A_2 E^n} [(\Psi^{AA'})^{\otimes n} \otimes \Phi^{A_1 B_1}], \quad (47)$$

$$\left\| \rho^{A^n B^n E^n} - (\Psi^{ABE})^{\otimes n} \right\|_1 \leq \text{error}. \quad (48)$$

The number  $(\log |B_2|)/n$  is called the **quantum communication rate** and  $(\log |B_1|)/n$  is called the **entanglement consumption rate** of the protocol.

A pair of real numbers  $(\mathcal{R}_Q, \mathcal{R}_E)$  is called an **achievable rate pair** if there exist, for  $n \rightarrow \infty$ , FQRS protocols with quantum communication rate approaching  $\mathcal{R}_Q$ , entanglement consumption rate approaching  $\mathcal{R}_E$ , and error approaching 0.

The achievable rates are described by the following theorem.

**Theorem 7** (Abeyesinghe *et al*, 2009 [20]). *The following rates are achievable for the FQRS:*

$$\mathcal{R}_Q > \frac{1}{2}I(A : B)_\Psi \quad \text{and} \quad \mathcal{R}_E < \mathcal{R}_Q + H(B|A)_\Psi. \quad (49)$$

We now provide the random coding exponents for the achievability of this protocol.

**Theorem 8.** *For any  $n \in \mathbb{N}$ , there exists a  $(\Psi, \text{error}, n)$  FQRS protocol for any  $\alpha \in (1, 2]$ ,  $\delta_1, \delta_2 > 0$ , such that*

$$\frac{\log |B_2|}{n} = \frac{1}{2} [H_{\tilde{\alpha}}(B)_\Psi - H_\alpha(B|A)_\Psi] + (|E| + 1)|A| \frac{\log(n+1)}{n} + \frac{\delta_1 + \delta_2}{2}, \quad (50)$$

$$\frac{\log |B_1|}{n} = \frac{\log |B_2|}{n} + H_\alpha(B|A)_\Psi - |A| \frac{\log(n+1)}{n} - \delta_2, \quad (51)$$

and the error approaches 0 exponentially in  $n$ .

*Proof.* We note the insightful observation in Refs. [29, 20] that FQRS can be implemented using ideas from FQSW.

Let  $W^{B^n \rightarrow B_1 B_2}$ ,  $|B_1| |B_2| \leq |B|^n$ , be a full-rank partial isometry. Then, using Corollary 2, we claim that there exists a Unitary  $U$  over  $B^n$  such that for  $\alpha \in (1, 2]$ ,

$$\begin{aligned} & \left\| \text{Tr}_{B_1 B_2} \circ \mathcal{T}_W^{B^n \rightarrow B_1 B_2} [U \cdot (\Psi^{ABE})^{\otimes n}] - (\Psi^{AE})^{\otimes n} \right\|_1 \\ & \leq 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |A| |E| \log(n+1) - n H_\alpha(B|AE)_\Psi + \Theta(\text{Tr}_{B_1 B_2} \circ \mathcal{T}_W) \right] \right\} \\ & = 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |A| |E| \log(n+1) + n H_{\tilde{\alpha}}(B)_\Psi - \log |B_1| |B_2| \right] \right\} \equiv \varepsilon_n, \end{aligned} \quad (52)$$

and

$$\begin{aligned} & \left\| \text{Tr}_{B_2} \circ \mathcal{T}_W^{B^n \rightarrow B_1 B_2} [U \cdot (\Psi^{AB})^{\otimes n}] - \pi^{B_1} \otimes (\Psi^A)^{\otimes n} \right\|_1 \\ & \leq 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |A| \log(n+1) - n H_\alpha(B|A)_\Psi + \log \frac{|B_1|}{|B_2|} \right] \right\} \equiv \vartheta_n. \end{aligned} \quad (53)$$

Using (53) and Lemma 31, we claim that there exists an isometry  $\tilde{U}^{B_2 E^n \rightarrow A_1 \tilde{B}^n \tilde{E}^n}$  such that

$$\left\| \tilde{U} \cdot \mathcal{T}_W^{B^n \rightarrow B_1 B_2} [U \cdot (\Psi^{ABE})^{\otimes n}] - \Phi^{A_1 B_1} \otimes (\Psi^{A \tilde{B} \tilde{E}})^{\otimes n} \right\|_1 \leq \Xi(\vartheta_n). \quad (54)$$

Using the compressive map  $\mathcal{C}_{\tilde{U}^\dagger} : A_1 \tilde{B}^n \tilde{E}^n \rightarrow A_2 E^n$ , (54), and monotonicity, we get

$$\left\| W^\dagger \cdot \mathcal{T}_W^{B^n \rightarrow B_1 A_2} [U \cdot (\Psi^{AB})^{\otimes n}] - W^\dagger \cdot \mathcal{C}_{\tilde{U}^\dagger} \left[ \Phi^{A_1 B_1} \otimes (\Psi^{A \tilde{B} \tilde{E}})^{\otimes n} \right] \right\|_1 \leq \Xi(\vartheta_n). \quad (55)$$

Using (52) and Lemma 31, we claim that there exists a Unitary  $V$  over  $B^n$  such that

$$\left\| W^\dagger \cdot \mathcal{T}_W^{B^n \rightarrow B_1 B_2} [U \cdot (\Psi^{ABE})^{\otimes n}] - V \cdot (\Psi^{ABE})^{\otimes n} \right\|_1 \leq \Xi(\varepsilon_n). \quad (56)$$

Using the triangle inequality, (55), and (56), we now have

$$\left\| (V^\dagger W^\dagger) \circ \mathcal{I}^{A_2 \rightarrow B_2} \circ \mathcal{C}_{\tilde{U}^\dagger} \left[ \Phi^{A_1 B_1} \otimes (\Psi^{A \tilde{B} \tilde{E}})^{\otimes n} \right] - (\Psi^{ABE})^{\otimes n} \right\|_1 \leq \Xi(\varepsilon_n) + \Xi(\vartheta_n). \quad (57)$$

Hence, the FQRS protocol consists of Alice applying

$$\mathcal{E}^{A'^n A_1 \rightarrow A_2 E^n} = \mathcal{C}_{\tilde{U}^\dagger}^{A_1 \tilde{B}^n \tilde{E}^n \rightarrow A_2 E^n} \circ \left( V_{\mathcal{N}}^{A' \rightarrow \tilde{B} \tilde{E}} \right)^{\otimes n}, \quad (58)$$

and Bob applying  $\mathcal{D}^{B_1 B_2 \rightarrow B^n} = V^\dagger W^\dagger$ . The claim now follows readily.  $\square$

## 7 Quantum state merging (QSM)

**Definition 6.** A  $(\Psi, \text{error}, n)$  QSM protocol consists of  $n$  copies of a pure state  $|\Psi\rangle^{ABR}$  shared between Alice ( $A$ ), Bob ( $B$ ), and reference ( $R$ ) inaccessible to both Alice and Bob, a MES  $\Phi^{A_0 B_0}$  shared between Alice ( $A_0$ ) and Bob ( $B_0$ ), and a local operation and classical communication (locc) quantum operation  $\mathcal{M} : A^n A_0 \otimes B^n B_0 \rightarrow A_1 \otimes B_1 \tilde{B}_2^n B_2^n$  such that for

$$\rho^{A_1 B_1 \tilde{B}_2^n B_2^n R^n} \equiv \mathcal{M} \left[ (\Psi^{ABR})^{\otimes n} \otimes \Phi^{A_0 B_0} \right], \quad (59)$$

$$\left\| \rho^{A_1 B_1 \tilde{B}_2^n B_2^n R^n} - \Phi^{A_1 B_1} \otimes \Psi^{\tilde{B}_2^n B_2^n R^n} \right\| \leq \text{error}, \quad (60)$$

where  $\Phi^{A_1 B_1}$  is a MES shared between Alice ( $A_1$ ) and Bob ( $B_1$ ). The number  $(\log |A_0| - \log |A_1|)/n$  is called the **entanglement rate** of the protocol. A real number  $\mathcal{R}_E$  is called an **achievable rate** if there exist, for  $n \rightarrow \infty$ , QSM protocols of rate approaching  $\mathcal{R}_E$  and error approaching 0.

The achievable rate is given in the next theorem.

**Theorem 9** (Horodecki et al, 2005 [30]). *The following rates are achievable:*

$$\mathcal{R}_E > H(A|B)_\Psi. \quad (61)$$

Furthermore, there exists a QSM protocol that achieves this merging cost using one-way locc with a classical communication cost of  $I(A : R)_\Psi$  per input copy.

We prove the following.

**Theorem 10.** *For any  $n \in \mathbb{N}$ , there exists a  $(\Psi, \text{error}, n)$  QSM protocol using one-way locc for arbitrary  $\delta_1, \delta_2 > 0$ ,  $\alpha \in (1, 2]$ , such that*

$$\frac{\log |A_0| - \log |A_1|}{n} = H_{\tilde{\alpha}}(A|B)_\Psi + |R| \frac{\log(n+1)}{n} + \delta_1, \quad (62)$$

and a classical communication cost of at most

$$H_{\tilde{\alpha}}(A)_\Psi - H_\alpha(A|R)_\Psi + \frac{(|B| + 1)|R| \log(n+1) + 2}{n} + \delta_1 + \delta_2, \quad (63)$$

with the error approaching 0 exponentially in  $n$ .



*Proof.* Our line of attack is similar to that in Ref. [30], Corollary 3.11 in Ref. [17], and Theorem 5.2 in Ref. [21].

Let  $W^{A^n \rightarrow E}$ ,  $|E| \leq |A|^n$ , be a full-rank partial isometry. Let  $\zeta \equiv \frac{|E||A_0|}{|A_1|}$ ,  $J \equiv \lceil \zeta \rceil$ , and let  $M_x^{EA_0 \rightarrow A_1}$ ,  $x = 1, \dots, J$ ,  $|A_1| \leq |A_0||E|$ , be a set of measurement operators such that  $\sum_{x=1}^J M_x^\dagger M_x = \mathbb{1}^{EA_0}$ , where each  $M_x$  (except possibly when  $x = J$ ) is a full-rank partial isometry.

For any orthonormal basis  $\{|x\rangle^X\}$ ,  $x = 1, \dots, J$ , we define

$$\mathcal{E}^{EA_0 \rightarrow XA_1}(\sigma^{EA_0}) \equiv \sum_{x=1}^J |x\rangle \langle x|^X \otimes (M_x \cdot \sigma^{EA_0}) \quad (64)$$

$$\omega^{XA_1} \equiv \mathcal{E}^{EA_0 \rightarrow XA_1} \circ \mathcal{T}_W^{A^n \rightarrow E}(\pi^{A^n A_0}). \quad (65)$$

We have

$$\omega^{XA_1} = \mathcal{E}^{EA_0 \rightarrow XA_1}(\pi^{EA_0}) \quad (66)$$

$$= \frac{1}{\zeta} \sum_{x=1}^J |x\rangle \langle x|^X \otimes \pi^{A_1} - |J\rangle \langle J|^X \otimes \frac{P^{A_1}}{|E||A_0|} \quad (67)$$

$$= \frac{J}{\zeta} \pi^{XA_1} - |J\rangle \langle J|^X \otimes \frac{P^{A_1}}{|E||A_0|}, \quad (68)$$

where  $P^{A_1}$  is a projector with rank  $< A_1$ . Note that

$$\|\omega^{XA_1} - \pi^{XA_1}\|_1 \leq \left\| \left( \frac{J}{\zeta} - 1 \right) \pi^{XA_1} \right\|_1 + \left\| \frac{P^{A_1}}{|E||A_0|} \right\|_1 < \left( \frac{J}{\zeta} - 1 \right) + \frac{1}{\zeta} < \frac{2}{\zeta}, \quad (69)$$

where the first inequality follows from the triangle inequality, the second one from  $\text{Tr} P^{A_1} < |A_1|$ , and the third one from  $J - \zeta < 1$ .

Invoking Corollary 2, we first claim that there exists a Unitary  $U^{A^n A_0}$  such that for any  $\alpha \in (1, 2]$ ,

$$\begin{aligned} & \left\| \mathcal{E}^{A^n A_0 \rightarrow A_1 X} \circ \mathcal{T}_W^{A^n \rightarrow E} \left\{ U^{A^n A_0} \cdot [(\Psi^{AR})^{\otimes n} \otimes \pi^{A_0}] \right\} - \omega^{A_1 X} \otimes (\Psi^R)^{\otimes n} \right\|_1 \\ & \leq 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |R| \log(n+1) - H_\alpha(A^n A_0 | R^n)_{\Psi^{\otimes n}} + \Theta(\mathcal{E} \circ \mathcal{T}_W) \right] \right\} \\ & \leq 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |R| \log(n+1) + n H_{\tilde{\alpha}}(A|B)_\Psi - (\log |A_0| - \log |A_1|) \right] \right\} \equiv \vartheta_n, \end{aligned} \quad (70)$$

(where in the second inequality, we have used  $-H_\alpha(A^n A_0 | R^n)_{\Psi^{\otimes n}} = -n H_\alpha(A|R)_\Psi - \log |A_0| = n H_{\tilde{\alpha}}(A|B)_\Psi - \log |A_0|$ , and, from Lemma 22,  $\Theta(\mathcal{E} \circ \mathcal{T}_W) \leq \log |A_1|$ ) and

$$\begin{aligned} & \left\| \text{Tr}_{EA_0} \circ \mathcal{T}_W^{A^n \rightarrow E} \left\{ U^{A^n A_0} \cdot [(\Psi^{ABR})^{\otimes n} \otimes \Phi^{A_0 B_0}] \right\} - (\Psi^{BR})^{\otimes n} \otimes \pi^{B_0} \right\|_1 \\ & \leq 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |B||R| \log(n+1) - H_\alpha(A^n A_0 | B^n R^n B_0)_{\Psi^{\otimes n} \otimes \Phi} - \log(|A_0||E|) \right] \right\} \\ & \leq 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |B||R| \log(n+1) + n H_{\tilde{\alpha}}(A)_\Psi - \log |E| \right] \right\} \equiv \varepsilon_n, \end{aligned} \quad (71)$$

where in the first inequality, we have used  $\nu_{(\Psi^{BR})^{\otimes n} \otimes \pi^{B_0}} = \nu_{(\Psi^{BR})^{\otimes n}}$ , and the second inequality follows from

$$\begin{aligned} & -H_\alpha(A^n A_0 | B^n R^n B_0)_{\Psi^{\otimes n} \otimes \Phi} - \log(|A_0||E|) \\ & \leq -H_\alpha(A^n | B^n R^n)_{\Psi^{\otimes n}} - H_\alpha(A_0 | B_0)_\Phi - \log(|A_0||E|) = nH_{\tilde{\alpha}}(A)_\Psi - \log|E|. \end{aligned} \quad (72)$$

(71) implies using Lemma 31 that there exists a Unitary  $V^{A^n A_0 \rightarrow A^n A_0}$  such that

$$\begin{aligned} & \|W^\dagger \cdot \mathcal{T}_W^{A^n \rightarrow E} \{U^{A^n A_0} \cdot [(\Psi^{ABR})^{\otimes n} \otimes \Phi^{A_0 B_0}]\} - V^{A^n A_0 \rightarrow A^n A_0} \cdot (\Psi^{ABR})^{\otimes n} \otimes \Phi^{A_0 B_0}\|_1 \\ & \leq \Xi(\varepsilon_n), \end{aligned} \quad (73)$$

and, using monotonicity, this implies that

$$\|\mathcal{T}_W^{A^n \rightarrow E} \{U^{A^n A_0} \cdot [(\Psi^{ABR})^{\otimes n} \otimes \Phi^{A_0 B_0}]\} - \mathcal{C}_W[V \cdot (\Psi^{ABR})^{\otimes n} \otimes \Phi^{A_0 B_0}]\|_1 \leq \Xi(\varepsilon_n). \quad (74)$$

We now have

$$\begin{aligned} & \|\mathcal{E}^{EA_0 \rightarrow A_1 X} \circ \mathcal{T}_W^{A^n \rightarrow E} [U \cdot (\Psi^{AR})^{\otimes n} \otimes \pi^{A_0}] - \pi^{A_1 X} \otimes (\Psi^R)^{\otimes n}\|_1 \\ & \leq \|\mathcal{E}^{EA_0 \rightarrow A_1 X} \circ \mathcal{T}_W^{A^n \rightarrow E} [U \cdot (\Psi^{AR})^{\otimes n} \otimes \pi^{A_0}] - \omega^{A_1 X} \otimes (\Psi^R)^{\otimes n}\|_1 \\ & \quad + \|\omega^{A_1 X} \otimes (\Psi^R)^{\otimes n} - \pi^{A_1 X} \otimes (\Psi^R)^{\otimes n}\|_1 \\ & \leq \vartheta_n + \frac{2}{\zeta} \equiv \beta_n, \end{aligned} \quad (75)$$

where the first inequality follows from the triangle inequality, and in the second inequality, the first term is upper bounded using (70), and the last term by using (69). Let

$$\xi_x^{A_1 B^n B_0 R^n} \equiv \frac{J|A|^n}{|E|} (M_x W U) \cdot (\Psi^{ABR})^{\otimes n} \otimes \Phi^{A_0 B_0} \quad (76)$$

$$\begin{aligned} \sigma^{XA_1 B^n B_0 R^n} & \equiv \mathcal{E}^{EA_0 \rightarrow A_1 X} \circ \mathcal{T}_W^{A^n \rightarrow E} [U \cdot (\Psi^{ABR})^{\otimes n} \otimes \Phi^{A_0 B_0}] \\ & = \sum_{x=1}^J \frac{1}{J} |x\rangle \langle x|^X \otimes \xi_x^{A_1 B^n B_0 R^n}. \end{aligned} \quad (77)$$

Note that  $\xi_x^{A_1 B^n B_0 R^n}$  is a pure state for all  $x$ . Let  $\varepsilon'_x \equiv \|\xi_x^{A_1 B^n} - \pi^{A_1} \otimes (\Psi^R)^{\otimes n}\|_1$ . We now have

$$\beta_n \geq \|\sigma^{XA_1 B^n} - \pi^{A_1 X} \otimes (\Psi^R)^{\otimes n}\|_1 = \sum_{x=1}^J \frac{\varepsilon'_x}{J}. \quad (78)$$

From Lemma 31, let  $V_x^{B^n B_0 \rightarrow B_1 \tilde{B}_2^n B_2^n}$  be an isometry such that

$$\left\| V_x^{B^n B_0 \rightarrow B_1 \tilde{B}_2^n B_2^n} \cdot \xi_x^{A_1 B^n B_0 R^n} - \Phi^{A_1 B_1} \otimes (\Psi^{\tilde{B}_2 B_2 R})^{\otimes n} \right\|_1 \leq \Xi(\varepsilon'_x). \quad (79)$$

Define a ctp map

$$\mathcal{D}^{XB^n B_0 \rightarrow B_1 \tilde{B}_2^n B_2^n} \left( \sum_{x=1}^J |x\rangle \langle x|^X \otimes \Upsilon_x^{B^n B_0} \right) \equiv \sum_{x=1}^J (V_x \cdot \Upsilon_x^{B^n B_0}). \quad (80)$$

We now have

$$\left\| \mathcal{D} \circ \mathcal{E} \circ \mathcal{T}_W [U \cdot (\Psi^{ABR})^{\otimes n} \otimes \Phi^{A_0 B_0}] - \Phi^{A_1 B_1} \otimes (\Psi^{\tilde{B}_2 B_2 R})^{\otimes n} \right\|_1 \quad (81)$$

$$= \left\| \sum_{x=1}^J \frac{1}{J} V_x^{B^n B_0 \rightarrow B_1 \tilde{B}_2^n B_2^n} \cdot \xi_x^{A_1 B^n B_0 R^n} - \Phi^{A_1 B_1} \otimes (\Psi^{\tilde{B}_2 B_2 R})^{\otimes n} \right\|_1 \quad (82)$$

$$\leq \sum_{x=1}^J \frac{1}{J} \left\| V_x^{B^n B_0 \rightarrow B_1 \tilde{B}_2^n B_2^n} \cdot \xi_x^{A_1 B^n B_0 R^n} - \Phi^{A_1 B_1} \otimes (\Psi^{\tilde{B}_2 B_2 R})^{\otimes n} \right\|_1 \quad (83)$$

$$\leq \sum_{x=1}^J \frac{1}{J} \Xi(\varepsilon'_x) \quad (84)$$

$$\leq \sum_{x=1}^J \frac{1}{J} \left[ 2\sqrt{\varepsilon'_x} + \sqrt{2}(\varepsilon'_x)^{3/4} + \varepsilon'_x \right] \quad (85)$$

$$\leq 2\sqrt{\sum_{x=1}^J \frac{\varepsilon'_x}{J}} + \sqrt{2} \left( \sum_{x=1}^J \frac{\varepsilon'_x}{J} \right)^{3/4} + \sum_{x=1}^J \frac{\varepsilon'_x}{J} \quad (86)$$

$$\leq 2\sqrt{\beta_n} + \sqrt{2}(\beta_n)^{3/4} + \beta_n, \quad (87)$$

where the first inequality follows from the convexity of the trace norm, the second inequality follows from (79), the third inequality follows since

$$\Xi(\varepsilon) = \sqrt{\varepsilon(2 + \varepsilon + 2\sqrt{1 + \varepsilon})} \leq 2\sqrt{\varepsilon} + \sqrt{2}(\varepsilon)^{3/4} + \varepsilon, \quad (88)$$

and the fourth inequality from the concavity of  $x \mapsto x^y$ ,  $y \in [0, 1]$ , and the last inequality follows from (78). Using (74) and the triangle inequality, we have

$$\begin{aligned} & \left\| \mathcal{D} \circ \mathcal{E} \circ \mathcal{C}_W [V \cdot (\Psi^{ABR})^{\otimes n} \otimes \Phi^{A_0 B_0}] - \Phi^{A_1 B_1} \otimes (\Psi^{\tilde{B}_2 B_2 R})^{\otimes n} \right\|_1 \\ & \leq \Xi(\varepsilon_n) + 2\sqrt{\beta_n} + \sqrt{2}(\beta_n)^{3/4} + \beta_n. \end{aligned} \quad (89)$$

Alice performs  $\mathcal{E}^{A^n A_0 \rightarrow A_1 X} \circ \mathcal{C}_W \circ V$  and Bob performs  $\mathcal{D}^{XB_0 B^n \rightarrow B_1 \tilde{B}_2^n B_2^n}$ . Note that  $J = \lceil \frac{|E||A_0|}{|A_1|} \rceil \leq \max\{1, \frac{2|E||A_0|}{|A_1|}\}$ , determines the classical communication cost. We have now shown the existence of a state merging protocol using one-way locc for arbitrary  $\delta_1, \delta_2 > 0$ ,

$\alpha \in (1, 2]$ , with

$$\frac{1}{n} \log \frac{|A_0|}{|A_1|} = H_{\tilde{\alpha}}(A|B)_{\Psi} + |R| \frac{\log(n+1)}{n} + \delta_1 \quad (90)$$

$$\frac{\log |E|}{n} = H_{\tilde{\alpha}}(A)_{\Psi} + |B||R| \frac{\log(n+1)}{n} + \delta_2 \quad (91)$$

$$\frac{\log J}{n} \leq H_{\tilde{\alpha}}(A)_{\Psi} - H_{\alpha}(A|R)_{\Psi} + \frac{(|B|+1)|R| \log(n+1) + 1}{n} + \delta_1 + \delta_2 \quad (92)$$

that has the error converging to 0 exponentially in  $n$ .  $\square$

## 8 Entanglement-assisted quantum communication with side information at the transmitter (Father with side information at the transmitter)

The definitions are directly from Ref. [17].

**Definition 7.** Let  $\mathcal{N}^{A'S \rightarrow B}$  be a cptp map with Stinespring dilation  $V_{\mathcal{N}}^{A'S \rightarrow BE}$  and  $|\Upsilon\rangle^{SS'}$  be a pure state. Then the transmitter encodes its information contained in  $\rho^{A_1 R} \in \mathcal{D}(\mathcal{H}_{A_1 R})$  using a cptp map  $\mathcal{E}^{A_1 S' \rightarrow A'}$ , and the output of the channel is  $\rho^{BR} = \mathcal{N}^{A'S \rightarrow B} \circ \mathcal{E}^{A_1 S' \rightarrow A'}(\rho^{A_1 R} \otimes \Upsilon^{SS'})$ . We denote this channel by  $\{\mathcal{N}^{A'S \rightarrow B}, |\Upsilon\rangle^{SS'}\}$ .

**Definition 8.** A  $(\{\mathcal{N}, |\Upsilon\rangle\}, \text{error}, n)$  father protocol with side information at the transmitter consists of  $n$  copies of two MES  $\Phi^{A_0 R}$  and  $\Phi^{A_1 B_1}$ , where Alice has  $A_0, A_1$ , Bob has  $B_1$ , and the reference  $R$  is inaccessible to both Alice and Bob, Alice applying an encoding map  $\mathcal{E}^{A_0 A_1 S'^n \rightarrow A'^n}$  to  $(\Phi^{A_0 R} \otimes \Phi^{A_1 B_1} \otimes \Upsilon^{SS'})^{\otimes n}$ ,  $n$  uses of the channel with side information at the transmitter  $\{\mathcal{N}^{A'S \rightarrow B}, |\Upsilon\rangle^{SS'}\}$ , and Bob applying a decoding map  $\mathcal{D}^{B^n B_1^n \rightarrow B_2^n}$  such that for

$$\rho^{B_2^n R^n} \equiv \mathcal{D}^{B^n B_1^n \rightarrow B_2^n} \circ (\mathcal{N}^{A'S \rightarrow B})^{\otimes n} \circ \mathcal{E}^{A_0 A_1 S'^n \rightarrow A'^n}(\Phi^{A_0 R} \otimes \Phi^{A_1 B_1} \otimes \Upsilon^{SS'})^{\otimes n}, \quad (93)$$

$$\|\rho^{B_2^n R^n} - (\Phi^{B_2 R})^{\otimes n}\|_1 \leq \text{error}. \quad (94)$$

The number  $\log |B_1|$  is called the **entanglement rate** of the protocol and  $\log |R|$  is called the **quantum communication rate** of the protocol.

A pair of real numbers  $(\mathcal{R}_Q, \mathcal{R}_E)$  is called an **achievable rate pair** if there exist, for  $n \rightarrow \infty$ , protocols with quantum communication rate approaching  $\mathcal{R}_Q$ , entanglement gain rate approaching  $\mathcal{R}_E$ , and error approaching 0.

The achievable rates are described by the following theorem.

**Theorem 11** (Dupuis, 2009 [17]). Let  $|\Psi\rangle^{CAA'S}$  be a pure state with  $\mathcal{H}_A = \mathcal{H}_R \otimes \mathcal{H}_{B_1}$  such that  $\Psi^S = \Upsilon^S$ , and  $|\Psi\rangle^{CABE} = V_{\mathcal{N}}^{A'S \rightarrow BE} |\Psi\rangle^{CAA'S}$ . The following rates are achievable:

$$\mathcal{R}_Q + \mathcal{R}_E < H(A|S)_{\Psi} \quad (95)$$

$$\mathcal{R}_Q - \mathcal{R}_E < -H(A|B)_{\Psi}. \quad (96)$$

We now have the following theorem.

**Theorem 12.** *For any  $n \in \mathbb{N}$ , and  $\Psi$  as defined in Theorem 11, there exists a  $(\{\mathcal{N}, |\Upsilon\rangle\}, \text{error}, n)$  Father protocol with side information at the transmitter such that for any  $\alpha \in (1, 2]$  and  $\delta_1, \delta_2 > 0$ ,*

$$\log |R| + \log |B_1| = H_\alpha(A|S)_\Psi - |S| \frac{\log(n+1)}{n} - \delta_1 \quad (97)$$

$$\log |R| - \log |B_1| = -H_{\tilde{\alpha}}(A|B)_\Psi - |C||E| \frac{\log(n+1)}{n} - \delta_2, \quad (98)$$

and the error approaches 0 exponentially in  $n$ .

*Proof.* We first claim using Corollary 2 that there exists a Unitary  $U$  on  $R^n B_1^n$  such that

$$\begin{aligned} & \left\| \text{Tr}_{B_1^n} [U \cdot (\Psi^{CRB_1 E})^{\otimes n}] - (\pi^R \otimes \Psi^{CE})^{\otimes n} \right\|_1 \\ & \leq 8 \exp \left\{ \frac{\alpha-1}{2\alpha} \left[ |C||E| \log(n+1) - nH_\alpha(A|CE)_\Psi + n \log \frac{|R|}{|B_1|} \right] \right\} \\ & = 8 \exp \left\{ \frac{\alpha-1}{2\alpha} \left[ |C||E| \log(n+1) + nH_{\tilde{\alpha}}(A|B)_\Psi + n \log \frac{|R|}{|B_1|} \right] \right\} \equiv \varepsilon_n, \end{aligned} \quad (99)$$

and

$$\begin{aligned} & \left\| U \cdot (\Psi^{RB_1 S})^{\otimes n} - (\pi^{RB_1})^{\otimes n} \otimes (\Upsilon^S)^{\otimes n} \right\|_1 = \left\| U \cdot (\Psi^{RB_1 S})^{\otimes n} - (\pi^{RB_1})^{\otimes n} \otimes (\Psi^S)^{\otimes n} \right\|_1 \\ & \leq 8 \exp \left\{ \frac{\alpha-1}{2\alpha} \left[ |S| \log(n+1) - nH_\alpha(A|S)_\Psi + n \log(|R||B_1|) \right] \right\} \equiv \vartheta_n, \end{aligned} \quad (100)$$

where in (100), we have used  $\Psi^S = \Upsilon^S$ , and it follows from (100) and Lemma 31 that there exists an isometry  $V_1^{A_0^n A_1^n S'^n \rightarrow A'^n C^n}$  such that

$$\left\| U \cdot (\Psi^{CRB_1 A' S})^{\otimes n} - V_1^{A_0^n A_1^n S'^n \rightarrow A'^n C^n} \cdot (\Phi^{A_0 R} \otimes \Phi^{A_1 B_1} \otimes \Upsilon^{S' S})^{\otimes n} \right\|_1 \leq 2\sqrt{\vartheta_n}. \quad (101)$$

Using the triangle inequality, (99), (101), and monotonicity, we have

$$\begin{aligned} & \left\| \text{Tr}_{B_1^n B^n} \left\{ \left[ (V_{\mathcal{N}})^{\otimes n} V_1^{A_0^n A_1^n S'^n \rightarrow A'^n C^n} \right] \cdot (\Phi^{A_0 R} \otimes \Phi^{A_1 B_1} \otimes \Upsilon^{S' S})^{\otimes n} \right\} - (\pi^R \otimes \Psi^{CE})^{\otimes n} \right\|_1 \\ & \leq \varepsilon_n + 2\sqrt{\vartheta_n}. \end{aligned} \quad (102)$$

Hence there exists an isometry  $V_2^{B_1^n B^n \rightarrow B_2^n \tilde{A}^n \tilde{B}^n}$  such that for some purifications  $\Phi^{RB_2}$  and  $\Psi^{\tilde{A}\tilde{B}CE}$  of  $\pi^R$  and  $\Psi^{CE}$  respectively, we have

$$\begin{aligned} & \left\| \left[ V_2^{B_1^n B^n \rightarrow B_2^n \tilde{A}^n \tilde{B}^n} (V_{\mathcal{N}})^{\otimes n} V_1^{A_0^n A_1^n \rightarrow A'^n} \right] \cdot (\Phi^{A_0 R} \otimes \Phi^{A_1 B_1} \otimes \Upsilon^{S' S})^{\otimes n} - (\Phi^{RB_2} \otimes \Psi^{\tilde{A}\tilde{B}CE})^{\otimes n} \right\|_1 \\ & \leq 2\sqrt{\varepsilon_n + 2\sqrt{\vartheta_n}} \end{aligned} \quad (103)$$

and hence,

$$\begin{aligned}
& \left\| \text{Tr}_{\tilde{A}^n \tilde{B}^n C^n} \left\{ V_2 \cdot (\mathcal{N})^{\otimes n} \left[ V_1 \cdot (\Phi^{A_0 R} \otimes \Phi^{A_1 B_1} \otimes \Upsilon^{S' S})^{\otimes n} \right] \right\} - (\Phi^{RB_2})^{\otimes n} \right\|_1 \\
&= \left\| \text{Tr}_{\tilde{A}^n \tilde{B}^n C^n E^n} \left\{ [V_2(V_{\mathcal{N}})^{\otimes n} V_1] \cdot (\Phi^{A_0 R} \otimes \Phi^{A_1 B_1} \otimes \Upsilon^{S' S})^{\otimes n} \right\} - (\Phi^{RB_2})^{\otimes n} \right\|_1 \\
&\leq 2\sqrt{\varepsilon_n} + 2\sqrt{\vartheta_n}.
\end{aligned} \tag{104}$$

It is now clear that Alice just applies  $\text{Tr}_{C^n} \circ V_1^{A_0^n A_1^n S'^m \rightarrow A'^n C^n}$  and Bob applies  $\text{Tr}_{\tilde{A}^n \tilde{B}^n} \circ V_2^{B_1^n B^n \rightarrow B_2^n \tilde{A}^n \tilde{B}^n}$ . The claim now follows readily.  $\square$

We now have the following corollary to obtain the regularized expressions by additional blocking.

**Corollary 13.** *For any  $m, n \in \mathbb{N}$ , a pure state  $\Psi^{CAA'^m S^m}$  with  $\mathcal{H}_A = \mathcal{H}_R \otimes \mathcal{H}_{B_1}$  such that  $\Psi^{S^m} = (\Upsilon^S)^{\otimes m}$  and  $|\Psi\rangle^{CAB^m E^m} = (V_{\mathcal{N}}^{A' S \rightarrow BE})^{\otimes m} |\Psi\rangle^{CAA'^m S^m}$ , there exists a  $(\{\mathcal{N}, |\Upsilon\rangle\}, \text{error}, mn)$  Fawther protocol with side information at the transmitter such that for any  $\alpha \in (1, 2]$  and  $\delta_1, \delta_2 > 0$ ,*

$$\frac{\log |R|}{m} + \frac{\log |B_1|}{m} = \frac{H_\alpha(A|S^m)_\Psi}{m} - |S| \frac{\log(mn+1)}{mn} - \delta_1 \tag{105}$$

$$\frac{\log |R|}{m} - \frac{\log |B_1|}{m} = -\frac{H_{\tilde{\alpha}}(A|B^m)_\Psi}{m} - |C||E|^m \frac{\log(n+1)}{mn} - \delta_2, \tag{106}$$

and the error approaches 0 exponentially in  $mn$ .

We omit the proof. Rather than blindly applying Theorem 12, we need to use  $\nu_{(\Upsilon^S)^{\otimes mn}} \leq (mn+1)^{|S|}$ . The number  $m$  serves two purposes. Firstly, it enables a better approximation to the optimal rates, and, secondly, it allows for finer approximation to the Rényi quantities through the choices of  $|R|$  and  $|B_1|$ .

Note that by choosing  $|B_1| = 1$ , we get **entanglement-unassisted quantum communication** as a special case of the above and for any  $\alpha \in (1, 2]$  and  $\delta_1, \delta_2 > 0$ , the rate is given by

$$\begin{aligned}
\frac{\log |R|}{m} = \min \Big\{ & \frac{H_\alpha(A|S^m)_\Psi}{m} - |S| \frac{\log(mn+1)}{mn} - \delta_1, \\
& -\frac{H_{\tilde{\alpha}}(A|B^m)_\Psi}{m} - |C||E|^m \frac{\log(n+1)}{mn} - \delta_2 \Big\}.
\end{aligned} \tag{107}$$

Assuming  $H_\alpha(A|S^m)_\Psi \geq -H_{\tilde{\alpha}}(A|B^m)_\Psi$ , the rate for **quantum communication assisted by unlimited entanglement** for any  $\alpha \in (1, 2]$  and  $\delta > 0$  is given by

$$\frac{\log |R|}{m} = \frac{H_\alpha(A|S^m)_\Psi - H_{\tilde{\alpha}}(A|B^m)_\Psi}{2m} - \frac{|S| \log(mn+1) + |C||E|^m \log(n+1)}{2mn} - \delta. \tag{108}$$



**Definition 9.** A  $(\{\mathcal{N}, |\Upsilon\rangle\}, \text{error}, n)$  **entanglement-assisted classical communication** protocol with side information at the transmitter consists of  $n$  copies of an MES  $\Phi^{A_2 B_2}$ , where Alice has  $A_2$  and Bob has  $B_2$ , Alice having a random variable  $X$  uniformly distributed over a set  $\mathcal{X}$  that models the information, Alice applying an encoding map  $\mathcal{E}_x^{A_2^{S'} \rightarrow A'^n}$ ,  $x \in \mathcal{X}$ , if  $X = x$ ,  $n$  uses of the channel with side information at the transmitter  $(\mathcal{N}^{A'S \rightarrow B}, |\Upsilon\rangle^{SS'})$ , and Bob applying a POVM (positive operator-valued measure)  $\{\Lambda_{x'}^{B^n B_2^n}, x' \in \mathcal{X}\}$ , such that for

$$\Pr\{x'|x\} \equiv \text{Tr} \Lambda_{x'}^{B^n B_2^n} \left[ (\mathcal{N}^{A'S \rightarrow B})^{\otimes n} \circ \mathcal{E}_x^{A_2^{S'} \rightarrow A'^n} (\Phi^{A_2 B_2} \otimes \Upsilon^{SS'})^{\otimes n} \right], \quad (109)$$

$$\frac{1}{|\mathcal{X}|} \sum_x (1 - \Pr\{x|x\}) \leq \text{error}. \quad (110)$$

The number  $(\log |\mathcal{X}|)/n$  is called the **classical communication rate** of the protocol.

A real number  $\mathcal{R}_C$  is called an **achievable rate** if there exist, for  $n \rightarrow \infty$ , any choice of  $|A_2|$ , protocols with classical communication rate approaching  $\mathcal{R}_C$  and error approaching 0.

Note that the capacity for this protocol was obtained in Ref. [17]. We now provide the random coding exponents for the entanglement-assisted classical communication.

**Corollary 14.** For any  $m, n \in \mathbb{N}$ , a pure state  $\Psi^{CAA^m S^m}$  with  $\mathcal{H}_A = \mathcal{H}_R \otimes \mathcal{H}_{B_1}$  such that  $\Psi^{S^m} = (\Upsilon^S)^{\otimes m}$  and  $|\Psi\rangle^{CAB^m E^m} = (V_{\mathcal{N}}^{A'S \rightarrow BE})^{\otimes m} |\Psi\rangle^{CAA^m S^m}$ , there exists a  $(\{\mathcal{N}, |\Upsilon\rangle\}, \text{error}, mn)$  **entanglement-assisted classical communication** protocol with side information at the transmitter such that for any  $\alpha \in (1, 2]$  and  $\delta > 0$ , the rate per channel use is given by

$$\frac{\log |\mathcal{X}|}{mn} = \frac{H_\alpha(A|S^m)_\Psi}{m} - \frac{H_{\tilde{\alpha}}(A|B^m)_\Psi}{m} - |S| \frac{\log(mn+1)}{mn} - |C||E|^m \frac{\log(n+1)}{mn} - \delta, \quad (111)$$

and the error approaches 0 exponentially in  $mn$ .

*Proof.* We follow the well-understood strategy to encapsulate the entanglement-assisted quantum communication protocol in the qudit superdense coding protocol. We follow the notation in Definition 8. Let  $\mathcal{H}_{A_2} = \mathcal{H}_{A_0} \otimes \mathcal{H}_{A_1}$  and  $\mathcal{H}_{B_2} = \mathcal{H}_R \otimes \mathcal{H}_{B_1}$ . Alice has access to  $A_0, A_1$  and Bob has access to  $R, B_1$ . Let  $V_i \in \mathbb{U}(R^n)$  such that  $\text{Tr} V_i^\dagger V_j = |R|^n \delta_{i,j}$ . Alice chooses  $|\mathcal{X}| = |R|^{2n}$ , and, for  $X = x$ , Alice applies  $V_x$  over  $R^n$  on  $(\Phi^{A_0 R})^{\otimes n}$  (Alice does this by exploiting the Schmidt symmetry) and passes that MES as input to the father protocol that uses the channel  $m \times n$  times. At the end of the father protocol, we have a state  $\rho_x^{B_1^{mn} R^{mn}}$  such that

$$\|\rho_x^{B_1^{mn} R^{mn}} - V_x \cdot (\Phi^{B_1 R})^{\otimes mn}\|_1 \leq \beta_{mn}, \quad (112)$$

where  $\beta_{mn} = 2\sqrt{\varepsilon_{m,n} + 2\sqrt{\vartheta_{m,n}}}$  and

$$\varepsilon_{m,n} = 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |C||E|^m \log(n+1) + n H_{\tilde{\alpha}}(A|B^m)_\Psi + n \log \frac{|R|}{|B_1|} \right] \right\}, \quad (113)$$

$$\vartheta_{m,n} = 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |S| \log(mn+1) - n H_\alpha(A|S^m)_\Psi + n \log(|R||B_1|) \right] \right\}, \quad (114)$$

and for appropriately chosen  $|R|$  and  $|B_1|$  as per (105) and (106),  $\beta_{mn}$  decays exponentially in  $mn$ . Bob now applies the POVM given by  $\{V_{x'} \cdot (\Phi^{B_2 R})^{\otimes mn}\}$ ,  $x' \in \mathcal{X}$ , and

$$\Pr\{x|x\} = \text{Tr} \rho_x^{B_2^{mn} R^{mn}} [V_x \cdot (\Phi^{B_2 R})^{\otimes mn}] = F [\rho_x^{B_2^{mn} R^{mn}}, V_x \cdot (\Phi^{B_2 R})^{\otimes mn}]^2 \geq 1 - \beta_{mn}, \quad (115)$$

where the inequality follows from the Fuchs-van de Graaf inequalities between trace distance and Fidelity [31] and in particular Corollary 9.3.2 in Ref. [5], and hence, the error of the protocol is upper bounded by  $\beta_{mn}$ . Lastly, it is easy to show that a ctp map followed by a POVM can be implemented just by a suitably chosen POVM, and hence, the decoder of the father protocol and the POVM of the superdense coding protocol can be implemented by a POVM. The claim now follows readily.  $\square$

## 9 Quantum state redistribution (QSR)

**Definition 10.** A  $(\Psi, \text{error}, n)$  QSR protocol consists of  $n$  copies of a pure state  $|\Psi\rangle^{ACBR}$  shared between with Alice ( $A$  and  $C$ ), Bob ( $B$ ), and the reference ( $R$ ) unavailable to both Alice and Bob, a MES  $\Phi^{A_1 B_1}$  shared between Alice ( $A_1$ ) and Bob ( $B_1$ ), Alice applying  $\mathcal{E} : A_1 C^n A^n \rightarrow C_2 C_3 \tilde{A}^n$ , a quantum communication across a noiseless quantum channel from Alice to Bob  $\mathcal{I}^{C_3 \rightarrow \tilde{B}}$ , and Bob applying  $\mathcal{D} : B_1 \tilde{B} B^n \rightarrow B_2 \tilde{B}_3^n B_3^n$  such that for

$$\rho^{C_2 B_2 \tilde{A}^n \tilde{B}_3^n B_3^n R^n} \equiv \mathcal{D}^{B_1 \tilde{B} B^n \rightarrow B_2 \tilde{B}_3^n B_3^n} \circ \mathcal{I}^{C_3 \rightarrow \tilde{B}} \circ \mathcal{E}^{A_1 C^n A^n \rightarrow C_2 C_3 \tilde{A}^n} [(\Psi^{ACBR})^{\otimes n} \otimes \Phi^{A_1 B_1}], \quad (116)$$

$$\left\| \rho^{C_2 B_2 \tilde{A}^n \tilde{B}_3^n B_3^n R^n} - \Phi^{C_2 B_2} \otimes (\Psi^{\tilde{A} \tilde{B}_3 B_3 R})^{\otimes n} \right\|_1 \leq \text{error}. \quad (117)$$

The number  $(\log |C_3|)/n$  is called the **quantum communication rate** and  $(\log |B_1| - \log |C_2|)/n$  is called the **entanglement cost rate** of the protocol.

A pair of real numbers  $(\mathcal{R}_Q, \mathcal{R}_E)$  is called an **achievable rate pair** if there exist, for  $n \rightarrow \infty$ , QSR protocols with quantum communication rate approaching  $\mathcal{R}_Q$ , entanglement cost rate approaching  $\mathcal{R}_E$ , and error approaching 0.

The achievable rates are described by the following theorem.

**Theorem 15** (Devetak and Yard, 2008 [32]). *The following rates are achievable for the QSR protocol:*

$$\mathcal{R}_Q > \frac{1}{2} I(C : R|B)_\Psi \quad \text{and} \quad \mathcal{R}_Q + \mathcal{R}_E > H(C|B)_\Psi. \quad (118)$$

Our goal in the remainder of this section is to provide the random coding exponents for the achievability of this protocol.

**Theorem 16.** *For any  $n \in \mathbb{N}$ , there exists a  $(\Psi, \text{error}, n)$  QSR protocol for any  $\alpha \in (1, 2]$ ,  $\delta_1, \delta_2 > 0$ , such that*

$$\frac{\log |C_3|}{n} = \frac{1}{2} [H_{\tilde{\alpha}}(C|B)_\Psi - H_\alpha(C|BR)_\Psi] + (|A| + |B|)|R| \frac{\log(n+1)}{2n} + \frac{\delta_1 + \delta_2}{2}, \quad (119)$$

$$\frac{1}{n} \log \frac{|C_3||B_1|}{|C_2|} = H_{\tilde{\alpha}}(C|B)_\Psi + |A||R| \frac{\log(n+1)}{n} + \delta_2, \quad (120)$$

and the error approaches 0 exponentially in  $n$ .

*Proof.* Our line of attack is similar to that in Ref. [33]. Let  $W^{C^n \rightarrow B_1 C_2 C_3}$ ,  $|B_1||C_2||C_3| \leq |C|^n$ , be a full-rank partial isometry. Then we can claim using Corollary 2 that for any  $\alpha \in (1, 2]$ , there exists a Unitary  $U^{C^n}$  such that

$$\begin{aligned} & \left\| \text{Tr}_{C_2 C_3} \circ \mathcal{T}_W^{C^n \rightarrow B_1 C_2 C_3} \left[ U^{C^n} \cdot (\Psi^{CBR})^{\otimes n} \right] - \pi^{B_1} \otimes (\Psi^{BR})^{\otimes n} \right\|_1 \\ & \leq 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |B||R| \log(n+1) - nH_\alpha(C|BR)_\Psi + \log \frac{|B_1|}{|C_2||C_3|} \right] \right\} \equiv \varepsilon_n, \end{aligned} \quad (121)$$

and

$$\begin{aligned} & \left\| \text{Tr}_{B_1 C_3} \circ \mathcal{T}_W^{C^n \rightarrow B_1 C_2 C_3} \left[ U^{C^n} \cdot (\Psi^{\tilde{A}CR})^{\otimes n} \right] - \pi^{C_2} \otimes (\Psi^{\tilde{A}R})^{\otimes n} \right\|_1 \\ & \leq 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |A||R| \log(n+1) - nH_\alpha(C|AR)_\Psi + \log \frac{|C_2|}{|B_1||C_3|} \right] \right\} \\ & = 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |A||R| \log(n+1) + nH_{\tilde{\alpha}}(C|B)_\Psi + \log \frac{|C_2|}{|B_1||C_3|} \right] \right\} \equiv \vartheta_n. \end{aligned} \quad (122)$$

Using (121), we claim that there exists an isometry  $V_1^{C_2 C_3 \tilde{A}^n \rightarrow A_1 A^n C^n}$  such that

$$\left\| V_1 \cdot \mathcal{T}_W \left[ U^{C^n} \cdot (\Psi^{\tilde{A}CBR})^{\otimes n} \right] - \Phi^{A_1 B_1} \otimes (\Psi^{ACBR})^{\otimes n} \right\|_1 \leq \Xi(\varepsilon_n), \quad (123)$$

and hence, using the compressive map  $\mathcal{C}_{V_1^\dagger} : A_1 A^n C^n \rightarrow C_2 C_3 \tilde{A}^n$ , we have

$$\left\| \mathcal{C}_{V_1^\dagger} \left[ \Phi^{A_1 B_1} \otimes (\Psi^{ACBR})^{\otimes n} \right] - \mathcal{T}_W \left[ U^{C^n} \cdot (\Psi^{\tilde{A}CBR})^{\otimes n} \right] \right\|_1 \leq \Xi(\varepsilon_n). \quad (124)$$

Using (122), we claim that there exists an isometry  $V_2^{B_1 \tilde{B} B^n \rightarrow B_2 \tilde{B}_3^n B_3^n}$  such that

$$\left\| V_2 \cdot \mathcal{I} \circ \mathcal{T}_W \left[ U^{C^n} \cdot (\Psi^{\tilde{A}CBR})^{\otimes n} \right] - \Phi^{C_2 B_2} \otimes (\Psi^{\tilde{A} \tilde{B}_3 B_3 R})^{\otimes n} \right\|_1 \leq \Xi(\vartheta_n). \quad (125)$$

Using monotonicity and triangle inequality, we now have

$$\left\| V_2 \cdot \mathcal{I} \circ \mathcal{C}_{V_1^\dagger} \left[ \Phi^{A_1 B_1} \otimes (\Psi^{ACBR})^{\otimes n} \right] - \Phi^{C_2 B_2} \otimes (\Psi^{\tilde{A} \tilde{B}_3 B_3 R})^{\otimes n} \right\|_1 \leq \Xi(\varepsilon_n) + \Xi(\vartheta_n). \quad (126)$$

Hence, Alice's operation is  $\mathcal{C}_{V_1^\dagger}^{A_1 A^n C^n \rightarrow C_2 C_3 \tilde{A}^n}$  and Bob's operation is  $V_2^{B_1 \tilde{B} B^n \rightarrow B_2 \tilde{B}_3^n B_3^n}$ . The claim now follows readily.  $\square$

## 10 Quantum communication across Broadcast Channels (QCBC)

**Definition 11.** A  $(\mathcal{N}, \text{error}, n)$  QCBC protocol consists of  $n$  copies of four MES  $|\Phi\rangle^{S_1 R_1}$ ,  $|\Phi\rangle^{A_1 B_1}$ ,  $|\Phi\rangle^{S_2 R_2}$ , and  $|\Phi\rangle^{A_2 B_2}$ , where Alice has  $S_1, S_2, A_1, A_2$ , Bob 1 has  $B_1$ , Bob 2 has  $B_2$ , and the references ( $R_1$  and  $R_2$ ) are inaccessible to both Alice and Bob, Alice applying the encoding map  $\mathcal{E}^{A_1 S_1 A_2 S_2 \rightarrow A'^n}$ ,  $n$  uses of a quantum broadcast channel from Alice to Bob 1 and 2,  $\mathcal{N}^{A' \rightarrow C_1 C_2}$  (with Stinespring dilation  $V_{\mathcal{N}}^{A' \rightarrow C_1 C_2 E}$ ), and local quantum operations by Bobs  $\mathcal{D}_i^{B_i C_i \rightarrow \tilde{S}_i^n}$ ,  $i = 1, 2$ , such that for

$$\rho^{\tilde{S}_1^n R_1^n \tilde{S}_2^n R_2^n} \equiv \left( \mathcal{D}_1^{B_1 C_1 \rightarrow \tilde{S}_1^n} \circ \mathcal{D}_2^{B_2 C_2 \rightarrow \tilde{S}_2^n} \right) \circ (\mathcal{N}^{A' \rightarrow C_1 C_2})^{\otimes n} \circ \mathcal{E}^{A_1 S_1 A_2 S_2 \rightarrow A'^n} \left[ (\Phi^{S_1 R_1} \otimes \Phi^{A_1 B_1} \otimes \Phi^{S_2 R_2} \otimes \Phi^{A_2 B_2})^{\otimes n} \right], \quad (127)$$

$$\left\| \rho^{\tilde{S}_1^n R_1^n \tilde{S}_2^n R_2^n} - (\Phi^{\tilde{S}_1 R_1} \otimes \Phi^{\tilde{S}_2 R_2})^{\otimes n} \right\|_1 \leq \text{error}. \quad (128)$$

For  $i = 1, 2$ , the numbers  $\log |R_i|$  are the **quantum communication rates** and  $\log |B_i|$  are the **entanglement consumption rates** of the protocol.

A vector of real numbers  $(\mathcal{R}_{Q,1}, \mathcal{R}_{Q,2}, \mathcal{R}_{E,1}, \mathcal{R}_{E,2})$  is called an **achievable rate vector** if there exist, for  $n \rightarrow \infty$ , QCBC protocols with quantum communication rates approaching  $\mathcal{R}_{Q,i}$ , entanglement consumption rates approaching  $\mathcal{R}_{E,i}$ ,  $i = 1, 2$ , and error approaching 0.

**Theorem 17** (Dupuis, 2009 [17]). Let  $|\Psi\rangle^{G_1 G_2 A' D}$  be any pure state with  $|\Psi\rangle^{G_1 G_2 C_1 C_2 E D} = V_{\mathcal{N}}^{A' \rightarrow C_1 C_2 E} |\Psi\rangle^{G_1 G_2 A' D}$ . The following rates are achievable:

$$\log |R_1| + \log |B_1| < H(G_1)_{\Psi} \quad (129)$$

$$\log |R_2| + \log |B_2| < H(G_2)_{\Psi} \quad (130)$$

$$\log |R_1| + \log |B_1| + \log |R_2| + \log |B_2| < H(G_1 G_2)_{\Psi} \quad (131)$$

$$\log |R_1| - \log |B_1| < I(G_1 C_1)_{\Psi} \quad (132)$$

$$\log |R_2| - \log |B_2| < I(G_2 C_2)_{\Psi}. \quad (133)$$

We follow the line of attack in Ref. [17] that we need to show the following theorem, which would yield Theorem 17. The regularized expressions can be obtained by additional blocking.

**Theorem 18.** For any  $n \in \mathbb{N}$ ,  $|\Psi\rangle^{G_1 G_2 A' D}$  and  $|\Psi\rangle^{G_1 G_2 C_1 C_2 E D}$  the states defined in Theorem 17, there exists a  $(\mathcal{N}, \text{error}, n)$  QCBC protocol such that for any  $\alpha \in (1, 2]$ ,  $\delta_1, \delta_2, \delta_3, \delta_4 > 0$ ,

$$\log |R_1| + \log |B_1| = H_{\alpha}(G_1 | G_2)_{\Psi} - \frac{|G_2| \log(n+1)}{n} - \delta_1 \quad (134)$$

$$\log |R_1| - \log |B_1| = -H_{\alpha}(G_1 | C_1)_{\Psi} - \frac{|G_2 C_2 E D| \log(n+1)}{n} - \delta_2 \quad (135)$$

$$\log |R_2| + \log |B_2| = H_{\alpha}(G_2)_{\Psi} - \delta_3 \quad (136)$$

$$\log |R_2| - \log |B_2| = -H_{\alpha}(G_2 | C_2)_{\Psi} - \frac{|G_1 C_1 E D| \log(n+1)}{n} - \delta_4 \quad (137)$$

and the error approaches 0 exponentially in  $n$ .

*Proof.* Let  $W_i^{G_i^n \rightarrow R_i^n B_i^n}$ ,  $|G_i|^n \geq |R_i|^n |B_i|^n$ ,  $i = 1, 2$ , be full-rank partial isometries. Define:

$$\varepsilon_{n,1} \equiv 20 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |G_2| \log(n+1) - nH_\alpha(G_1|G_2)_\Psi + n \log |R_1| |B_1| \right] \right\} \quad (138)$$

$$\varepsilon_{n,2} \equiv 20 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |G_2 C_2 E D| \log(n+1) - nH_\alpha(G_1|G_2 C_2 E D)_\Psi + n \log \frac{|R_1|}{|B_1|} \right] \right\} \quad (139)$$

$$\varepsilon_{n,3} \equiv 20 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ -nH_\alpha(G_2)_\Psi + n \log(|R_2| |B_2|) \right] \right\} \quad (140)$$

$$\varepsilon_{n,4} \equiv 20 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ |G_1 C_1 E D| \log(n+1) - nH_\alpha(G_2|G_1 C_1 E D)_\Psi + n \log \frac{|R_2|}{|B_2|} \right] \right\} \quad (141)$$

$$\varepsilon_{n,5} \equiv 20 \exp \left\{ \frac{\alpha - 1}{2\alpha} \left[ -nH_\alpha(G_2)_\Psi - n \log |G_2| \right] \right\}. \quad (142)$$

For  $i = 1, 2$ , let  $U_i$  be random Unitaries on  $G_i^n$ . We have

$$\begin{aligned} & \mathbb{E}_{U_1, U_2} \left\| \mathcal{T}_{W_1}^{G_1^n \rightarrow R_1^n B_1^n} \circ \mathcal{T}_{W_2}^{G_2^n \rightarrow R_2^n B_2^n} [(U_1 \otimes U_2) \cdot (\Psi^{G_1 G_2})^{\otimes n}] - (\pi^{R_1 B_1 R_2 B_2})^{\otimes n} \right\|_1 \\ & \leq \mathbb{E}_{U_1, U_2} \left\| \mathcal{T}_{W_2}^{G_2^n \rightarrow R_2^n B_2^n} \left( U_2 \cdot \left\{ \mathcal{T}_{W_1}^{G_1^n \rightarrow R_1^n B_1^n} [U_1 \cdot (\Psi^{G_1 G_2})^{\otimes n}] - (\pi^{R_1 B_1})^{\otimes n} \otimes (\Psi^{G_2})^{\otimes n} \right\} \right) \right\|_1 \\ & \quad + \mathbb{E}_{U_2} \left\| (\pi^{R_1 B_1})^{\otimes n} \otimes \mathcal{T}_{W_2}^{G_2^n \rightarrow R_2^n B_2^n} [U_2 \cdot (\Psi^{G_2})^{\otimes n}] - (\pi^{R_1 B_1 R_2 B_2})^{\otimes n} \right\|_1 \\ & \leq \mathbb{E}_{U_1} \left\| \mathcal{T}_{W_1}^{G_1^n \rightarrow R_1^n B_1^n} [U_1 \cdot (\Psi^{G_1 G_2})^{\otimes n}] - (\pi^{R_1 B_1})^{\otimes n} \otimes (\Psi^{G_2})^{\otimes n} \right\|_1 \\ & \quad + \mathbb{E}_{U_2} \left\| \mathcal{T}_{W_2}^{G_2^n \rightarrow R_2^n B_2^n} [U_2 \cdot (\Psi^{G_2})^{\otimes n}] - (\pi^{R_2 B_2})^{\otimes n} \right\|_1 \\ & \leq (\varepsilon_{n,1} + \varepsilon_{n,3})/5, \end{aligned} \quad (143)$$

where the first inequality follows from the triangle's inequality and the second inequality follows since  $\mathcal{T}_{W_2}$  is a class-1 map and the last inequality from Theorem 1. We also have

$$\begin{aligned} & \mathbb{E}_{U_1, U_2} \left\| \mathcal{T}_{W_2} (U_2 \cdot \{ \text{Tr}_{B_1} \circ \mathcal{T}_{W_1} [U_1 \cdot (\Psi^{G_1 G_2 C_2 E D})^{\otimes n}] - (\pi^{R_1} \otimes \Psi^{G_2 C_2 E D})^{\otimes n} \}) \right\|_1 \\ & \leq \mathbb{E}_{U_1} \left\| \text{Tr}_{B_1} \circ \mathcal{T}_{W_1} [U_1 \cdot (\Psi^{G_1 G_2 C_2 E D})^{\otimes n}] - (\pi^{R_1} \otimes \Psi^{G_2 C_2 E D})^{\otimes n} \right\|_1 \leq \varepsilon_{n,2}/5, \end{aligned} \quad (144)$$

where the first inequality follows since  $\mathcal{T}_{W_2}$  is a class-1 map,

$$\mathbb{E}_{U_2} \left\| \text{Tr}_{B_2} \circ \mathcal{T}_{W_2} [U_2 \cdot (\Psi^{G_1 G_2 C_1 E D})^{\otimes n}] - (\pi^{R_2} \otimes \Psi^{G_1 C_1 E D})^{\otimes n} \right\|_1 \leq \varepsilon_{n,4}/5, \quad (145)$$

$$\mathbb{E}_{U_2} \left| \text{Tr} \circ \mathcal{T}_{W_2} [U_2 \cdot (\Psi^{G_2})^{\otimes n}] - 1 \right| \leq \varepsilon_{n,5}/5. \quad (146)$$

We now use the arguments in Corollary 2 to claim that there exist Unitaries  $U_i$  on  $G_i^n$ ,  $i = 1, 2$ , such that

$$\left\| \mathcal{T}_{W_1} \circ \mathcal{T}_{W_2} [(U_1 \otimes U_2) \cdot (\Psi^{G_1 G_2})^{\otimes n}] - (\pi^{R_1 B_1 R_2 B_2})^{\otimes n} \right\|_1 \leq \varepsilon_{n,1} + \varepsilon_{n,3} \quad (147)$$

$$\left\| \mathcal{T}_{W_2} (U_2 \cdot \{ \text{Tr}_{B_1} \circ \mathcal{T}_{W_1} [U_1 \cdot (\Psi^{G_1 G_2 C_2 E D})^{\otimes n}] - (\pi^{R_1} \otimes \Psi^{G_2 C_2 E D})^{\otimes n} \}) \right\|_1 \leq \varepsilon_{n,2} \quad (148)$$

$$\left\| \text{Tr}_{B_2} \circ \mathcal{T}_{W_2} [U_2 \cdot (\Psi^{G_1 G_2 C_1 E D})^{\otimes n}] - (\pi^{R_2} \otimes \Psi^{G_1 C_1 E D})^{\otimes n} \right\|_1 \leq \varepsilon_{n,4} \quad (149)$$

$$\left| \text{Tr} \circ \mathcal{T}_{W_2} [U_2 \cdot (\Psi^{G_2})^{\otimes n}] - 1 \right| \leq \varepsilon_{n,5}. \quad (150)$$

It now follows that there exist isometries  $V_1^{S_1^n A_1^n S_2^n A_2^n \rightarrow A'^n D^n}$ ,  $V_2^{B_1^n C_1^n \rightarrow \tilde{G}_1^n \tilde{S}_1^n \tilde{C}_1^n}$ , and  $V_3^{B_2^n C_2^n \rightarrow \tilde{G}_2^n \tilde{S}_2^n \tilde{C}_2^n}$  such that

$$\left\| \mathcal{T}_{W_1} \circ \mathcal{T}_{W_2} \left[ (U_1 \otimes U_2) \cdot (\Psi^{G_1 G_2 A' D})^{\otimes n} \right] - V_1 \cdot (\Phi^{S_1 R_1} \otimes \Phi^{A_1 B_1} \otimes \Phi^{S_2 R_2} \otimes \Phi^{A_2 B_2})^{\otimes n} \right\|_1 \leq \Xi(\varepsilon_{n,1} + \varepsilon_{n,3}), \quad (151)$$

$$\left\| \mathcal{T}_{W_2} \left( U_2 \cdot \left\{ V_2 \circ \mathcal{T}_{W_1} \left[ U_1 \cdot (\Psi^{G_1 G_2 C_1 C_2 E D})^{\otimes n} \right] - (\Phi^{R_1 \tilde{S}_1})^{\otimes n} \otimes (\Psi^{\tilde{G}_1 \tilde{G}_2 \tilde{C}_1 \tilde{C}_2 E D})^{\otimes n} \right\} \right) \right\|_1 \leq \Xi(\varepsilon_{n,2}) + \varepsilon_{n,5}, \quad (152)$$

where we have used the triangle's inequality, and

$$\left\| V_3 \cdot \mathcal{T}_{W_2}^{G_2^n \rightarrow R_2^n B_2^n} \left[ U_2 \cdot (\Psi^{G_1 G_2 C_1 C_2 E D})^{\otimes n} \right] - (\Phi^{R_2 \tilde{S}_2})^{\otimes n} \otimes (\Psi^{G_1 \tilde{G}_2 C_1 \tilde{C}_2 E D})^{\otimes n} \right\|_1 \leq \Xi(\varepsilon_{n,4}). \quad (153)$$

We now have for

$$\mathcal{E}^{A_1 S_1 A_2 S_2 \rightarrow A'^n} \equiv \text{Tr}_{D^n} \circ V_1^{S_1^n A_1^n S_2^n A_2^n \rightarrow A'^n D^n}, \quad (154)$$

$$\mathcal{D}_1^{B_1^n C_1^n \rightarrow \tilde{S}_1^n} \equiv \text{Tr}_{\tilde{G}_1^n \tilde{C}_1^n} \circ V_2^{B_1^n C_1^n \rightarrow \tilde{G}_1^n \tilde{S}_1^n \tilde{C}_1^n}, \quad (155)$$

$$\mathcal{D}_2^{B_2^n C_2^n \rightarrow \tilde{S}_2^n} \equiv \text{Tr}_{\tilde{G}_2^n \tilde{C}_2^n} \circ V_3^{B_2^n C_2^n \rightarrow \tilde{G}_2^n \tilde{S}_2^n \tilde{C}_2^n}, \quad (156)$$

$$\Upsilon^{R_1^n \tilde{S}_1^n R_2^n \tilde{S}_2^n} \equiv \mathcal{D}_1 \circ \mathcal{D}_2 \circ \mathcal{T}_{W_1} \circ \mathcal{T}_{W_2} \left[ (U_1 \otimes U_2) \cdot (\Psi^{G_1 G_2 C_1 C_2})^{\otimes n} \right], \quad (157)$$

$$\Upsilon_2^{R_1^n \tilde{S}_1^n R_2^n \tilde{S}_2^n} \equiv (\Phi^{R_1 \tilde{S}_1})^{\otimes n} \otimes \mathcal{D}_2 \circ \mathcal{T}_{W_2} \left[ U_2 \cdot (\Psi^{G_2 C_2})^{\otimes n} \right], \quad (158)$$

$$\begin{aligned} & \left\| \mathcal{D}_1 \circ \mathcal{D}_2 \circ (\mathcal{N})^{\otimes n} \circ \mathcal{E} \left[ (\Phi^{S_1 R_1} \otimes \Phi^{A_1 B_1} \otimes \Phi^{S_2 R_2} \otimes \Phi^{A_2 B_2})^{\otimes n} \right] - (\Phi^{R_1 \tilde{S}_1} \otimes \Phi^{R_2 \tilde{S}_2})^{\otimes n} \right\|_1 \\ & \leq \left\| \Upsilon^{R_1^n \tilde{S}_1^n R_2^n \tilde{S}_2^n} - (\Phi^{R_1 \tilde{S}_1} \otimes \Phi^{R_2 \tilde{S}_2})^{\otimes n} \right\|_1 + \Xi(\varepsilon_{n,1} + \varepsilon_{n,3}) \\ & \leq \left\| \Upsilon^{R_1^n \tilde{S}_1^n R_2^n \tilde{S}_2^n} - \Upsilon_2^{R_1^n \tilde{S}_1^n R_2^n \tilde{S}_2^n} \right\|_1 + \left\| \Upsilon_2^{R_1^n \tilde{S}_1^n R_2^n \tilde{S}_2^n} - (\Phi^{R_1 \tilde{S}_1} \otimes \Phi^{R_2 \tilde{S}_2})^{\otimes n} \right\|_1 + \Xi(\varepsilon_{n,1} + \varepsilon_{n,3}) \\ & \leq \Xi(\varepsilon_{n,2}) + \varepsilon_{n,5} + \Xi(\varepsilon_{n,4}) + \Xi(\varepsilon_{n,1} + \varepsilon_{n,3}), \end{aligned} \quad (159)$$

where the first inequality follows from (151), the triangle inequality and the monotonicity, the second inequality follows from the triangle inequality, the third inequality follows from (152), (153), and monotonicity. The claim of the Theorem now follows from (159).  $\square$

**Remark:** It is clear from the above theorem that any rate in the following rate region is achievable with error decaying exponentially in  $n$  to zero:

$$\log |R_1| + \log |B_1| < H(G_1 | G_2)_\Psi \quad (160)$$

$$\log |R_2| + \log |B_2| < H(G_2)_\Psi \quad (161)$$

$$\log |R_1| - \log |B_1| < I(G_1 | C_1)_\Psi \quad (162)$$

$$\log |R_2| - \log |B_2| < I(G_2 | C_2)_\Psi \quad (163)$$



We now repeat the argument given in Theorem 5.3 in Ref. [17] that by switching the roles of Bob 1 and Bob 2 and doing time sharing, we can achieve any point in the rate region as stipulated in the claim of Theorem 17.

## 11 Destroying correlations by adding classical randomness

**Definition 12.** A  $(\rho, \text{error}, n)$  protocol for destroying correlations by adding classical randomness consists of  $n$  copies of a bipartite state  $\rho^{AR}$ , and applying  $M$  Unitaries  $U_i$ ,  $i = 1, \dots, M$ , over  $A^n$  such that

$$\left\| \frac{1}{M} \sum_{i=1}^M [U_i \cdot (\rho^{AR})^{\otimes n}] - \sigma^{A^n} \otimes (\rho^R)^{\otimes n} \right\|_1 \leq \text{error}, \quad (164)$$

where  $\sigma^{A^n} \in \mathcal{D}(\mathcal{H}_{A^n})$  and we make no apriori restrictions on the choice of  $\sigma^{A^n}$ .

The number  $(\log M)/n$  is called the **rate** of the protocol. A real numbers  $\mathcal{R}_C$  is called an **achievable rate** if there exist, for  $n \rightarrow \infty$ , protocols with rate approaching  $\mathcal{R}_C$  and the error approaching 0.

**Theorem 19** (Groisman et al, 2005 [34]). The smallest achievable rate is  $I(A : R)_\rho$ .

We prove the following theorem.

**Theorem 20.** For any  $n \in \mathbb{N}$ , there exists a  $(\rho, \text{error}, n)$  protocol such that for any  $\delta > 0$ ,  $\alpha \in (1, 2]$  and  $|\Psi\rangle^{ARE}$  a purification of  $\rho^{AR}$ ,

$$\frac{\log M}{n} = H_{\tilde{\alpha}}(A)_\rho - H_\alpha(A|R)_\rho + (|E| + 1)|R| \frac{\log(n+1)}{n} + \delta, \quad (165)$$

and the error approaches 0 exponentially in  $n$ .

*Proof.* Consider a partial isometry  $W^{A^n \rightarrow B}$ ,  $|B| \leq |A^n|$ . For  $M \leq |B|^2$ , we can choose  $M$  Unitaries  $V_i^B \in \mathbb{U}(B)$  such that  $\text{Tr}(V_i^B)^\dagger V_j^B = |B| \delta_{i,j}$ , and let  $\mathcal{V}_M : B \rightarrow B$  be a cptp map given by

$$\mathcal{V}_M(\sigma^B) \equiv \frac{1}{M} \sum_{i=1}^M V_i^B \cdot \sigma^B. \quad (166)$$

Then, from Corollary 2, for any  $\alpha \in (1, 2]$ , there exists a Unitary  $U$  such that

$$\begin{aligned} & \left\| \text{Tr}_B \circ \mathcal{T}_W [U \cdot (\Psi^{ARE})^{\otimes n}] - (\Psi^{RE})^{\otimes n} \right\|_1 \\ & \leq 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} [|R||E| \log(n+1) + n H_{\tilde{\alpha}}(A)_\rho - \log |B|] \right\} \equiv \varepsilon_n, \end{aligned} \quad (167)$$

and

$$\begin{aligned} & \left\| \mathcal{V}_M \circ \mathcal{T}_W [U \cdot (\rho^{AR})^{\otimes n}] - \pi^B \otimes (\rho^R)^{\otimes n} \right\|_1 \\ & \leq 8 \exp \left\{ \frac{\alpha - 1}{2\alpha} [|R| \log(n+1) - n H_\alpha(A|R)_\rho - \log M + \log |B|] \right\} \equiv \vartheta_n, \end{aligned} \quad (168)$$

where we have used  $\Theta(\mathcal{V}_M \circ \mathcal{T}_W) \leq \log |B| - \log M$  from Lemma 23. From (167) and Lemma 31, we claim that there exists a Unitary  $U_2$  over  $A^n$  such that

$$\|W^\dagger \cdot \mathcal{T}_W [U \cdot (\Psi^{ARE})^{\otimes n}] - U_2 \cdot (\Psi^{ARE})^{\otimes n}\|_1 \leq \Xi(\varepsilon_n). \quad (169)$$

Consider now the following Unitaries over  $A^n$  constructed from  $V_i^B$  as  $V_i^{A^n} = W^\dagger \cdot V_i^B + (\mathbb{1}^A - W^\dagger W)$ . Note that  $V_i^{A^n} W^\dagger = W^\dagger V_i^B$ . We now claim that  $V_i^{A^n} U_2$  are the  $M$  Unitaries we need. We have

$$\begin{aligned} & \left\| \frac{1}{M} \sum_{i=1}^M (V_i^{A^n} U_2) \cdot (\rho^{AR})^{\otimes n} - (W^\dagger \cdot \pi^B) \otimes (\rho^R)^{\otimes n} \right\|_1 \\ & \leq \left\| \frac{1}{M} \sum_{i=1}^M (V_i^{A^n} U_2) \cdot (\rho^{AR})^{\otimes n} - \frac{1}{M} \sum_{i=1}^M (V_i^{A^n} W^\dagger) \cdot \mathcal{T}_W [U \cdot (\rho^{AR})^{\otimes n}] \right\|_1 + \\ & \quad \left\| \frac{1}{M} \sum_{i=1}^M (V_i^{A^n} W^\dagger) \cdot \mathcal{T}_W [U \cdot (\rho^{AR})^{\otimes n}] - (W^\dagger \cdot \pi^B) \otimes (\rho^R)^{\otimes n} \right\|_1 \end{aligned} \quad (170)$$

$$\begin{aligned} & \leq \frac{1}{M} \sum_{i=1}^M \left\| (V_i^{A^n} U_2) \cdot (\rho^{AR})^{\otimes n} - (V_i^{A^n} W^\dagger) \cdot \mathcal{T}_W [U \cdot (\rho^{AR})^{\otimes n}] \right\|_1 + \\ & \quad \left\| \frac{1}{M} \sum_{i=1}^M (W^\dagger V_i^B) \cdot \mathcal{T}_W [U \cdot (\rho^{AR})^{\otimes n}] - (W^\dagger \cdot \pi^B) \otimes (\rho^R)^{\otimes n} \right\|_1 \end{aligned} \quad (171)$$

$$\begin{aligned} & \leq \frac{1}{M} \sum_{i=1}^M \left\| U_2 \cdot (\rho^{AR})^{\otimes n} - W^\dagger \cdot \mathcal{T}_W [U \cdot (\rho^{AR})^{\otimes n}] \right\|_1 + \\ & \quad \left\| \frac{1}{M} \sum_{i=1}^M V_i^B \cdot \mathcal{T}_W [U \cdot (\rho^{AR})^{\otimes n}] - \pi^B \otimes (\rho^R)^{\otimes n} \right\|_1 \end{aligned} \quad (172)$$

$$\leq \Xi(\varepsilon_n) + \vartheta_n, \quad (173)$$

where the first inequality follows from the triangle inequality, in the second inequality, the first term follows from the convexity of the trace norm and the second term follows by invoking  $V_i^{A^n} W^\dagger = W^\dagger V_i^B$ , in the third inequality, the first term follows by invoking the Unitary invariance of the trace norm and the second term from monotonicity, in the fourth inequality, the first term is upper bounded using (169) and the second term is upper bounded using (168). The claim now follows readily.  $\square$

## 12 Conclusions

In conclusion, we have provided a new version of the decoupling theorem that gives an exponential bound on the average decoupling error with a Rényi  $\alpha$ -conditional entropy in the exponent for a restricted class of completely positive maps for any  $\alpha \in (1, 2]$  as

opposed to only  $\alpha = 2$  in Ref. [17]. This key step allows us to make a connection with the random coding exponents, which we provide for several important protocols including those at the top of the family tree of protocols. The importance of random coding exponents for the achievability of information-processing tasks has been well known since the seminal work by Gallager [8]. Such an analysis, with very few exceptions thus far, has been missing and we now fill that void with this paper. The version of the decoupling theorem and other ideas developed in this paper may well find wider applications with or without further extensions.

## A Computation of $\Theta$ for some cases

**Lemma 21.** *For a full-rank partial isometry  $W^{A \rightarrow A_1 A_2}$ ,  $|A_1||A_2| \leq |A|$ ,*

$$\Theta(\text{Tr}_{A_2} \circ \mathcal{T}_W^{A \rightarrow A_1 A_2}) \leq \log \frac{|A_1|}{|A_2|} \quad (174)$$

$$\Theta(\text{Tr}_{A_2} \circ \mathcal{C}_W^{A \rightarrow A_1 A_2}) \leq \log \frac{|A_1|}{|A_2|}. \quad (175)$$

*Proof.* Since we have the freedom in choosing the local orthonormal bases in describing the MES, hence, let them be such that  $W|i\rangle^A = |i\rangle^{A_1 A_2}$  for  $i \leq |A_1||A_2|$ , and  $W|i\rangle^A = 0$  for  $i > |A_1||A_2|$ , where  $\{|i\rangle^A\}$  and  $\{|i\rangle^{A_1 A_2}\}$  are orthonormal states in their respective systems. It now follows that

$$\mathcal{T}_W(|i\rangle \langle j|^A) = \frac{|A|}{|A_1||A_2|} |i\rangle \langle j|^{A_1 A_2} \text{ind}_{\{i,j \leq |A_1||A_2|\}} \quad (176)$$

$$\mathcal{C}_W(|i\rangle \langle j|^A) = |i\rangle \langle j|^{A_1 A_2} \text{ind}_{\{i,j \leq |A_1||A_2|\}} + \delta_{i,j} \pi^{A_1 A_2} \text{ind}_{\{i,j > |A_1||A_2|\}}. \quad (177)$$

We now have

$$\exp\{\Theta(\text{Tr}_{A_2} \circ \mathcal{C}_W)\} \leq \frac{|A_1|}{|A|^2} \sum_{i,j} \text{Tr} \left[ \text{Tr}_{A_2} \circ \mathcal{C}_W(|i\rangle \langle j|^A) \right] \left[ \text{Tr}_{A_2} \circ \mathcal{C}_W(|j\rangle \langle i|^A) \right] \quad (178)$$

$$= \frac{|A_1|}{|A|^2} \text{Tr} \left[ \sum_{i,j \leq |A_1||A_2|} \text{Tr}_{A_2}(|i\rangle \langle j|^{A_1 A_2}) \text{Tr}_{A_2}(|j\rangle \langle i|^{A_1 A_2}) + \sum_{i,j > |A_1||A_2|} \delta_{i,j} (\text{Tr}_{A_2} \pi^{A_1 A_2})^2 \right] \quad (179)$$

$$= \frac{|A_1|}{|A|^2} \left( |A_1|^2 |A_2| + \frac{|A| - |A_1||A_2|}{|A_1|} \right) \leq \frac{|A_1|}{|A_2|}, \quad (180)$$

where the first inequality follows using (1). Following the above, we arrive at

$$\exp\{\Theta(\text{Tr}_{A_2} \circ \mathcal{T}_W)\} \leq \frac{|A_1|}{|A|^2} \left[ \left( \frac{|A|}{|A_1||A_2|} \right)^2 |A_1|^2 |A_2| \right] = \frac{|A_1|}{|A_2|}. \quad (181)$$

QED. □

**Lemma 22.** Let  $\{M_i \in \mathcal{L}(BC, D), i = 1, \dots, J\}$ ,  $J = \lceil \frac{BC}{D} \rceil$ , be a complete set of measurement operators ( $\sum_i M_i^\dagger M_i = \mathbb{1}^{BC}$ ). Let  $\zeta \equiv \frac{|B||C|}{|D|}$  and let the first  $\vartheta \equiv \lfloor \frac{BC}{D} \rfloor$   $M_i$ 's be rank- $|D|$  partial isometries. Define for any orthonormal basis  $\{|i\rangle^X\}$ ,  $i = 1, \dots, J$ ,

$$\mathcal{E}^{BC \rightarrow XD}(\sigma^{BC}) = \sum_{i=1}^J |i\rangle \langle i|^X \otimes (M_i \cdot \sigma^{BC}) \quad (182)$$

and let  $W^{A \rightarrow B}$ ,  $|B| \leq |A|$ , be a full-rank partial isometry. Then

$$\Theta(\mathcal{E} \circ \mathcal{T}_W) \leq \log |D|. \quad (183)$$

*Proof.* Let  $W^{A \rightarrow B} = \sum_{i=1}^{|B|} |i\rangle^B \langle i|^A$ . Once again, we exploit the freedom in choosing the local bases in defining MES and have

$$|\Phi\rangle^{AA'CC'} = \frac{1}{\sqrt{|A||C|}} \sum_{i_1, i_2} |i_1\rangle^A |i_1\rangle^{A'} |i_2\rangle^C |i_2\rangle^{C'}. \quad (184)$$

Hence,

$$X_{i_1, j_1} \equiv \mathcal{T}_W(|i_1\rangle \langle j_1|^A) = \frac{|A|}{|B|} |i_1\rangle \langle j_1|^B \text{ind}_{\{i_1, j_1 \leq |B|\}}. \quad (185)$$

We now have for  $\theta^{XD} = \sum_x p_x |x\rangle \langle x|^X \otimes \pi^D$ ,  $\{p_x\}$  a probability vector (whose choice is specified below),

$$\begin{aligned} & \exp\{\Theta(\mathcal{E} \circ \mathcal{T}_W)\} \\ & \leq \frac{1}{|A|^2 |C|^2} \sum_{i_1, j_1, i_2, j_2} \text{Tr} \left[ \mathcal{E}(X_{i_1, j_1} \otimes |i_2\rangle \langle j_2|^C) \mathcal{E}(X_{j_1, i_1} \otimes |j_2\rangle \langle i_2|^C) \right] (\theta^{XD})^{-1} \end{aligned} \quad (186)$$

$$\begin{aligned} & = \frac{1}{|B|^2 |C|^2} \sum_{i_1, j_1, i_2, j_2, x} \text{Tr} \left[ |x\rangle \langle x|^X \otimes M_x(|i_1\rangle \langle j_1|^B \otimes |i_2\rangle \langle j_2|^C) M_x^\dagger M_x \right. \\ & \quad \left. (|j_1\rangle \langle i_1|^B \otimes |j_2\rangle \langle i_2|^C) M_x^\dagger \right] (\theta^{XD})^{-1} \end{aligned} \quad (187)$$

$$= \frac{1}{|B|^2 |C|^2} \sum_{i_1, i_2, x} (\text{Tr} M_x M_x^\dagger) \text{Tr} \left[ |x\rangle \langle x|^X \otimes M_x(|i_1\rangle \langle i_1|^B \otimes |i_2\rangle \langle i_2|^C) M_x^\dagger \right] (\theta^{XD})^{-1} \quad (188)$$

$$= \frac{1}{|B|^2 |C|^2} \sum_x (\text{Tr} M_x M_x^\dagger)^2 \frac{|D|}{p_x}. \quad (189)$$

Let  $p = 1 - \vartheta/\zeta$ ,  $p_x = (1 - p)/\vartheta$  for  $x = 1, \dots, \vartheta$ , and if  $|D|$  doesn't divide  $|B||C|$ , then there

is an additional entry  $p_x = p$  if  $x = \vartheta + 1$ . Continuing from above, we now have

$$\exp\{\Theta(\mathcal{E} \circ \mathcal{T}_W)\} \leq \frac{1}{|B|^2|C|^2} \left[ \vartheta|D|^2 \frac{|D|}{\frac{1-p}{\vartheta}} + (|B||C| - |D|\vartheta)^2 \frac{|D|}{p} \right] \quad (190)$$

$$= |D| \left[ \frac{\vartheta^2}{\zeta^2(1-p)} + \left(1 - \frac{\vartheta}{\zeta}\right)^2 \frac{1}{p} \right] \quad (191)$$

$$= |D| \left[ \frac{\vartheta}{\zeta} + 1 - \frac{\vartheta}{\zeta} \right] = |D|. \quad (192)$$

QED.  $\square$

**Lemma 23.** For  $M \in \mathbb{N}$ ,  $M \leq |B|^2$ ,  $M$  Unitaries  $V_i^B \in \mathbb{U}(B)$  such that  $\text{Tr}(V_i^B)^\dagger V_j^B = |B|\delta_{i,j}$ , let  $\mathcal{V}_M : B \rightarrow B$  be a cptp map given by

$$\mathcal{V}_M(\sigma^B) \equiv \frac{1}{M} \sum_{i=1}^M V_i^B \cdot \sigma^B. \quad (193)$$

Then

$$\Theta(\mathcal{V}_M \circ \mathcal{T}_W) \leq \log |B| - \log M. \quad (194)$$

*Proof.* Let  $W|i\rangle^A = |i\rangle^B$  for  $i \leq |B|$ , and  $W|i\rangle^A = 0$  for  $i > |B|$ , where  $\{|i\rangle^A\}$  and  $\{|i\rangle^B\}$  are orthonormal states in their respective systems. Using  $\mathcal{T}_W(|i\rangle\langle j|^A) = \frac{|A|}{|B|} |i\rangle\langle j|^B \text{ind}_{\{i,j \leq |B|\}}$ , we have

$$\exp\{\Theta(\mathcal{V}_M \circ \mathcal{T}_W)\} \leq \frac{|B|}{|A|^2} \sum_{i,j} \text{Tr} \left[ \mathcal{V} \circ \mathcal{T}_W(|i\rangle\langle j|^A) \right] \left[ \mathcal{V} \circ \mathcal{T}_W(|j\rangle\langle i|^A) \right] \quad (195)$$

$$= \frac{1}{|B|} \text{Tr} \left[ \sum_{i,j \leq |B|} \mathcal{V}(|i\rangle\langle j|^B) \mathcal{V}(|j\rangle\langle i|^B) \right] \quad (196)$$

$$= \frac{1}{|B|M^2} \text{Tr} \left[ \sum_{i,j \leq |B|} \sum_{k,l=1}^M V_k |i\rangle\langle j|^B V_k^\dagger V_l |j\rangle\langle i|^B V_l^\dagger \right] \quad (197)$$

$$= \frac{1}{|B|M^2} \sum_{k,l=1}^M \left| \text{Tr} V_l^\dagger V_k \right|^2 = \frac{1}{|B|M^2} \sum_{k,l=1}^M |B|^2 \delta_{k,l} = \frac{|B|}{M}, \quad (198)$$

where the first inequality follows using (1) and the fourth equality follows since  $\text{Tr} V_l^\dagger V_k = |B|\delta_{k,l}$ . QED.  $\square$

## B Lemmata

**Lemma 24.** Let  $\mathcal{T}$  be a completely positive map. Then for any inputs  $\sigma, \theta$  (not necessarily Hermitian), there exists a contraction  $K$  such that

$$\mathcal{T}(\sigma\theta^\dagger)\mathcal{T}(\theta\sigma^\dagger) = \sqrt{\mathcal{T}(\sigma\sigma^\dagger)} K \mathcal{T}(\theta\theta^\dagger) K^\dagger \sqrt{\mathcal{T}(\sigma\sigma^\dagger)}. \quad (199)$$

In particular, if  $\theta = \mathbb{1}$  and  $\mathcal{T}(\mathbb{1})$  is a scaled identity, i.e., commutes with all matrices, then

$$\mathcal{T}(\sigma)\mathcal{T}(\sigma^\dagger) \leq \mathcal{T}(\sigma\sigma^\dagger)\mathcal{T}(\mathbb{1}). \quad (200)$$

An example of such a  $\mathcal{T}$  is the partial trace.

*Proof.* Since  $\mathcal{T}$  is completely positive, it is also 2-positive. Hence, if  $\mathcal{I}_2$  is the identity super-operator for  $2 \times 2$  matrices, then for orthonormal  $|0\rangle, |1\rangle$ , we have

$$0 \leq (\mathcal{I}_2 \otimes \mathcal{T}) [(|0\rangle \otimes \theta + |1\rangle \otimes \sigma)(|0\rangle \otimes \theta + |1\rangle \otimes \sigma)^\dagger] \quad (201)$$

$$= |0\rangle\langle 0| \otimes \mathcal{T}(\theta\theta^\dagger) + |1\rangle\langle 0| \otimes \mathcal{T}(\sigma\theta^\dagger) + |0\rangle\langle 1| \otimes \mathcal{T}(\theta\sigma^\dagger) + |1\rangle\langle 1| \otimes \mathcal{T}(\sigma\sigma^\dagger). \quad (202)$$

We now invoke Theorem IX.5.9 in Ref. [35] to claim that there exists a contraction  $K$  such that

$$\mathcal{T}(\sigma\theta^\dagger) = \sqrt{\mathcal{T}(\sigma\sigma^\dagger)}K\sqrt{\mathcal{T}(\theta\theta^\dagger)}. \quad (203)$$

The claim and the particular case now follow easily.  $\square$

**Lemma 25.** Let  $\mathcal{T}^{A \rightarrow E}$  be any completely positive map such that  $\text{Tr } \mathcal{T}(\mathbb{1}^A) = |A|$ . Then  $\mathcal{T}^{A \rightarrow E}$  is a class-1 map. For any ctp map  $\mathcal{E}^{E \rightarrow C}$ ,  $\mathcal{E}^{E \rightarrow C} \circ \mathcal{T}^{A \rightarrow E}$  is also a class-1 map.

*Proof.* Let the Kraus operators of  $\mathcal{T}$  be given by  $\{E_i\}$ . We have for a random Unitary  $U$  over  $A$  and any  $\sigma \in \mathcal{L}(\mathcal{H}_A)$ ,

$$\mathbb{E}_U \|\mathcal{T}(U \cdot \sigma)\|_1 = \frac{1}{|A|^2} \sum_j \|\mathcal{T}(U_j \cdot \sigma)\|_1 \quad (204)$$

$$= \left\| \frac{1}{|A|^2} \sum_{i,j} (|j\rangle^B \otimes E_i U_j) \cdot \sigma \right\|_1 \quad (205)$$

$$= \|\mathcal{F}(\sigma)\|_1 \quad (206)$$

$$\leq \|\sigma\|_1, \quad (207)$$

where in the second equality,  $\{|j\rangle^B\}$  is an orthonormal basis in  $B$ ,  $\mathcal{F}^{A \rightarrow BE}$  is a ctp map with Kraus operators  $\{\frac{1}{|A|}(|j\rangle^B \otimes E_i U_j)\}$ , and the last inequality is well known. The second statement of the claim follows simply by noting that  $\text{Tr } \mathcal{E} \circ \mathcal{T}(\mathbb{1}^A) = \text{Tr } \mathcal{T}(\mathbb{1}^A) = |A|$ . QED.  $\square$

**Lemma 26.** For any matrices  $\sigma^{AR}$ ,  $X^A$ ,  $W^R$  (not necessarily Hermitian) and for  $U$  acting on  $A$ , we have

$$\begin{aligned} \mathbb{E}_U \left\{ U \sigma^{AR} U^\dagger (X^A \otimes W^R) U (\sigma^{AR})^\dagger U^\dagger \right\} \\ = \frac{X^A \otimes (|A|\Lambda^R - \Upsilon^R) + (\text{Tr } X^A) \mathbb{1}^A \otimes (|A|\Upsilon^R - \Lambda^R)}{|A|(|A|^2 - 1)}, \end{aligned} \quad (208)$$

where  $\Lambda^R \equiv \sigma^R W^R (\sigma^R)^\dagger$  and  $\Upsilon^R \equiv \text{Tr}_A [\sigma^{AR} (\mathbb{1}^A \otimes W^R) (\sigma^{AR})^\dagger]$ .



*Proof.* Consider first vectors  $\{|\varphi_i\rangle\}$ ,  $i \in 1, \dots, 6$ , in  $\mathcal{H}_A$  and we have

$$\mathbb{E}_U \left\{ U |\varphi_1\rangle \langle \varphi_2| U^\dagger |\varphi_3\rangle \langle \varphi_4| U |\varphi_5\rangle \langle \varphi_6| U^\dagger \right\} \quad (209)$$

$$= (\mathbb{1} \otimes \langle \varphi_4|) \mathbb{E}_U \left\{ (U \otimes U)(|\varphi_1\rangle \langle \varphi_5|)(\langle \varphi_6| \langle \varphi_2|)(U^\dagger \otimes U^\dagger) \right\} (\mathbb{1} \otimes |\varphi_3\rangle) \quad (210)$$

$$= (\mathbb{1} \otimes \langle \varphi_4|) \left( \frac{q_1|A| - q_2}{|A|(|A|^2 - 1)} \mathbb{1}^{AA'} + \frac{q_2|A| - q_1}{|A|(|A|^2 - 1)} F^{AA'} \right) (\mathbb{1} \otimes |\varphi_3\rangle) \quad (211)$$

$$= \frac{q_1|A| - q_2}{|A|(|A|^2 - 1)} \langle \varphi_4| \varphi_3 \rangle \mathbb{1}^A + \frac{q_2|A| - q_1}{|A|(|A|^2 - 1)} |\varphi_3\rangle \langle \varphi_4|, \quad (212)$$

where the integral in the second equality is well known (see Lemma 3.4 in Ref. [17]),  $q_1 = \langle \varphi_6| \varphi_1 \rangle \langle \varphi_2| \varphi_5 \rangle$ ,  $q_2 = \langle \varphi_2| \varphi_1 \rangle \langle \varphi_6| \varphi_5 \rangle$ , and  $F^{AA'}$  is the swap operator. We have by singular value decomposition:

$$X^A = \sum_i \eta_i |y_i\rangle \langle z_i|^A. \quad (213)$$

We also have by the singular value and Schmidt decompositions:

$$\sigma^{AR} = \sum_{i,j,k} \sqrt{\beta_i^2 \lambda_{i,j} \mu_{i,k}} |v_{ij}\rangle \langle w_{ik}|^A \otimes |v_{ij}\rangle \langle w_{ik}|^R. \quad (214)$$

Let  $i_1^2 = (i_1, i_2)$ ,  $i_1^3 = (i_1, \dots, i_3)$ ,  $j_1^2 = (j_1, j_2)$  and  $k_1^2 = (k_1, k_2)$ . We now have

$$\begin{aligned} & \mathbb{E}_U \left\{ U \sigma^{AR} U^\dagger (X^A \otimes W^R) U (\sigma^{AR})^\dagger U^\dagger \right\} \\ &= \sum_{i_1^3, j_1^2, k_1^2} f_1(i_1^3, j_1^2, k_1^2) \mathbb{E}_U \left\{ U(|v_{i_1 j_1}\rangle \langle w_{i_1 k_1}|^A \otimes |v_{i_1 j_1}\rangle \langle w_{i_1 k_1}|^R) U^\dagger(|y_{i_3}\rangle \langle z_{i_3}|^A \otimes W^R) \right. \\ & \quad \left. U(|w_{i_2 j_2}\rangle \langle v_{i_2 k_2}|^A \otimes |w_{i_2 j_2}\rangle \langle v_{i_2 k_2}|^R) U^\dagger \right\} \end{aligned} \quad (215)$$

$$\begin{aligned} &= \sum_{i_1^3, j_1^2, k_1^2} f_1(i_1^3, j_1^2, k_1^2) \langle w_{i_1 k_1} | W | w_{i_2 j_2} \rangle^R \times \\ & \quad \mathbb{E}_U \left\{ U |v_{i_1 j_1}\rangle^A \langle w_{i_1 k_1} | U^\dagger |y_{i_3}\rangle^A \langle z_{i_3} | U |w_{i_2 j_2}\rangle^A \langle v_{i_2 k_2}|^A U^\dagger \right\} \otimes |v_{i_1 j_1}\rangle \langle v_{i_2 k_2}|^R \end{aligned} \quad (216)$$

$$\begin{aligned} &= \sum_{i_1^3, j_1^2, k_1^2} f_1(i_1^3, j_1^2, k_1^2) \langle w_{i_1 k_1} | W | w_{i_2 j_2} \rangle^R \left[ \frac{q_1(i_1^2, j_1^2, k_1^2)|A| - q_2(i_1^2, j_1^2, k_1^2)}{|A|(|A|^2 - 1)} \langle z_{i_3} | y_{i_3} \rangle^A \mathbb{1}^A + \right. \\ & \quad \left. \frac{q_2(i_1^2, j_1^2, k_1^2)|A| - q_1(i_1^2, j_1^2, k_1^2)}{|A|(|A|^2 - 1)} |y_{i_3}\rangle \langle z_{i_3}|^A \right] \otimes |v_{i_1 j_1}\rangle \langle v_{i_2 k_2}|^R \end{aligned} \quad (217)$$

$$= \frac{X^A \otimes (|A|\Lambda^R - \Upsilon^R) + (\text{Tr} X^A) \mathbb{1}^A \otimes (|A|\Upsilon^R - \Lambda^R)}{|A|(|A|^2 - 1)}, \quad (218)$$

where in the first equality

$$f_1(i_1^3, j_1^2, k_1^2) = \sqrt{\beta_{i_1}^2 \lambda_{i_1, j_1} \mu_{i_1, k_1} \eta_{i_3}^2 \beta_{i_2}^2 \lambda_{i_2, j_2} \mu_{i_2, k_2}}, \quad (219)$$

in the third equality,

$$q_1(i_1^2, j_1^2, k_1^2) = \langle w_{i_1 k_1} | w_{i_2 j_2} \rangle^A \langle v_{i_2 k_2} | v_{i_1 j_1} \rangle^A \quad (220)$$

$$q_2(i_1^2, j_1^2, k_1^2) = \langle w_{i_1 k_1} | v_{i_1 j_1} \rangle^A \langle v_{i_2 k_2} | w_{i_2 j_2} \rangle^A, \quad (221)$$

and the fourth equality follows after simplifications. QED.  $\square$

**Lemma 27.** Let  $\mathcal{T}^{A \rightarrow E}$  be a completely positive map with the Choi-Jamiołkowski representation  $\omega_{\mathcal{T}}^{EA'}$ . Then for a random Unitary  $U$  acting on  $A$ , any matrix  $\sigma^{AR}$ , we have

$$\begin{aligned} \mathbb{E}_U \left\{ [\mathcal{T}(U \cdot \sigma^{AR}) - \omega_{\mathcal{T}}^E \otimes \sigma^R] [\mathcal{T}(U \cdot \sigma^{AR}) - \omega_{\mathcal{T}}^E \otimes \sigma^R]^\dagger \right\} &= \frac{\mathcal{Q}_{A'}(\omega_{\mathcal{T}}^{EA'}) \otimes \mathcal{Q}_A(\sigma^{AR})}{|A|^2 - 1} \\ &\leq \frac{|A|^2}{|A|^2 - 1} \text{Tr}_{A'} \left( \omega_{\mathcal{T}}^{EA'} \right)^2 \otimes \text{Tr}_A [\sigma^{AR}(\sigma^{AR})^\dagger]. \end{aligned} \quad (222)$$

*Proof.* Let  $\mathcal{T}$  be described by the Kraus operators  $\{T_k\}$ . We now have

$$\begin{aligned} \mathbb{E}_U \left\{ [\mathcal{T}(U \cdot \sigma^{AR}) - \omega_{\mathcal{T}}^E \otimes \sigma^R] [\mathcal{T}(U \cdot \sigma^{AR}) - \omega_{\mathcal{T}}^E \otimes \sigma^R]^\dagger \right\} \\ = \sum_{k,l} T_k \mathbb{E}_U \left\{ U \sigma^{AR} U^\dagger T_k^\dagger T_l U (\sigma^{AR})^\dagger U^\dagger \right\} T_l^\dagger - (\omega_{\mathcal{T}}^E)^2 \otimes \sigma^R (\sigma^R)^\dagger \end{aligned} \quad (223)$$

$$\begin{aligned} = \sum_{k,l} T_k \left\{ T_k^\dagger T_l \otimes \frac{|A| \sigma^R (\sigma^R)^\dagger - \text{Tr}_A [\sigma^{AR} (\sigma^{AR})^\dagger]}{|A|(|A|^2 - 1)} + \right. \\ \left. (\text{Tr} T_k^\dagger T_l) \mathbb{1}^A \otimes \frac{|A| \text{Tr}_A [\sigma^{AR} (\sigma^{AR})^\dagger] - \sigma^R (\sigma^R)^\dagger}{|A|(|A|^2 - 1)} \right\} T_l^\dagger - (\omega_{\mathcal{T}}^E)^2 \otimes \sigma^R (\sigma^R)^\dagger \end{aligned} \quad (224)$$

$$\begin{aligned} = |A|^2 (\omega_{\mathcal{T}}^E)^2 \otimes \frac{|A| \sigma^R (\sigma^R)^\dagger - \text{Tr}_A [\sigma^{AR} (\sigma^{AR})^\dagger]}{|A|(|A|^2 - 1)} + \\ |A|^2 \text{Tr}_{A'} \left( \omega_{\mathcal{T}}^{EA'} \right)^2 \otimes \frac{|A| \text{Tr}_A [\sigma^{AR} (\sigma^{AR})^\dagger] - \sigma^R (\sigma^R)^\dagger}{|A|(|A|^2 - 1)} - (\omega_{\mathcal{T}}^E)^2 \otimes \sigma^R (\sigma^R)^\dagger \end{aligned} \quad (225)$$

$$\begin{aligned} = (\omega_{\mathcal{T}}^E)^2 \otimes \frac{|A|^2 \sigma^R (\sigma^R)^\dagger - |A| \text{Tr}_A [\sigma^{AR} (\sigma^{AR})^\dagger]}{|A|^2 - 1} + \\ |A| \text{Tr}_{A'} \left( \omega_{\mathcal{T}}^{EA'} \right)^2 \otimes \frac{|A| \text{Tr}_A [\sigma^{AR} (\sigma^{AR})^\dagger] - \sigma^R (\sigma^R)^\dagger}{|A|^2 - 1} - (\omega_{\mathcal{T}}^E)^2 \otimes \sigma^R (\sigma^R)^\dagger \end{aligned} \quad (226)$$

$$= \frac{\mathcal{Q}_{A'}(\omega_{\mathcal{T}}^{EA'}) \otimes \mathcal{Q}_A(\sigma^{AR})}{|A|^2 - 1} \quad (227)$$

$$\leq \frac{|A|^2}{|A|^2 - 1} \text{Tr}_{A'} \left( \omega_{\mathcal{T}}^{EA'} \right)^2 \otimes \text{Tr}_A [\sigma^{AR} (\sigma^{AR})^\dagger], \quad (228)$$

where in the second equality, we have used Lemma 26, and the inequality follows by noting from Lemma 24 that  $|A| \text{Tr}_A [\sigma^{AR} (\sigma^{AR})^\dagger] - \sigma^R (\sigma^R)^\dagger \in \text{Pos}(\mathcal{H}_R)$ . QED.  $\square$

**Lemma 28** (Exercise 9.9 in Ref. [3]). Let  $\rho \in D(\mathcal{H}_A)$ ,  $\sigma \in \text{Pos}(\mathcal{H}_A)$ , and  $\Pi = \{\mathcal{M}_\sigma(\rho) \geq \zeta\sigma\}$ . Then for any  $\alpha \in (1, 2]$ , we have

$$\|\Pi\rho\|_1 \leq \zeta^{\frac{1-\alpha}{2}} \sqrt{Q_\alpha(\rho\|\sigma)} = \zeta^{\frac{1-\alpha}{2}} \exp\left\{\frac{\alpha-1}{2} D_\alpha(\rho\|\sigma)\right\}. \quad (229)$$

**Lemma 29** (Hayashi [3]). Let  $\rho \in D(\mathcal{H}_A)$ ,  $\sigma \in \text{Pos}(\mathcal{H}_A)$ ,  $\Pi = \{\mathcal{M}_\sigma(\rho) \geq \zeta\sigma\}$  and  $\hat{\Pi} = \mathbb{1} - \Pi$ . Then

$$\text{Tr}\sigma^{-1}\hat{\Pi}\rho^2\hat{\Pi} \leq \nu_\sigma\zeta. \quad (230)$$

The proof of this lemma is contained in Lemma 9.2 in Ref. [3].

**Lemma 30.** Let  $\sigma, \rho \in \text{Pos}(\mathcal{H}_A)$ . Then

$$\text{Tr}\rho + \text{Tr}\sigma - 2F(\rho, \sigma) \leq \|\rho - \sigma\|_1 \leq \sqrt{(\text{Tr}\rho + \text{Tr}\sigma)^2 - 4F(\rho, \sigma)^2}. \quad (231)$$

*Proof.* The proof is essentially along the lines of the Fuchs-van de Graaf inequalities [31]. We know that

$$F(\rho, \sigma) = \min_{\text{POVM}\{\Lambda_m\}} \sum_m \sqrt{p_m q_m}, \quad (232)$$

where  $p_m \equiv \text{Tr}\Lambda_m\rho$  and  $q_m \equiv \text{Tr}\Lambda_m\sigma$ . Note that  $\sum_m p_m = \text{Tr}\rho$  and  $\sum_m q_m = \text{Tr}\sigma$ . Let  $\{\Lambda_m\}$  be the minimizing POVM in the above equation. We now have

$$\begin{aligned} \|\rho - \sigma\|_1 &\geq \left\| \sum_m |m\rangle \langle m|^X \otimes \sqrt{\Lambda_m}\rho\sqrt{\Lambda_m} - \sum_m |m\rangle \langle m|^X \otimes \sqrt{\Lambda_m}\sigma\sqrt{\Lambda_m} \right\|_1 \\ &\geq \left\| \sum_m p_m |m\rangle \langle m|^X - \sum_m q_m |m\rangle \langle m|^X \right\|_1 = \sum_m |p_m - q_m| \\ &= \sum_m |\sqrt{p_m} - \sqrt{q_m}| |\sqrt{p_m} + \sqrt{q_m}| \geq \sum_m (\sqrt{p_m} - \sqrt{q_m})^2 \\ &= \text{Tr}\rho + \text{Tr}\sigma - 2F(\rho, \sigma), \end{aligned} \quad (233)$$

where the first inequality follows from the monotonicity under the application of a cptp map with Kraus operators  $\{|m\rangle^X \otimes \sqrt{\Lambda_m}\}$ , where  $\{|m\rangle^X\}$  is an orthonormal basis, and the second inequality follows again from monotonicity under partial trace.

To prove the other inequality, let  $|u_\rho\rangle$  and  $|v_\sigma\rangle$  be purifications of  $\rho$  and  $\sigma$  respectively such that  $F(\rho, \sigma) = \langle u_\rho | v_\sigma \rangle$ . We now have

$$\|\rho - \sigma\|_1 \leq \|u_\rho - v_\sigma\|_1 = \sqrt{(\text{Tr}\rho + \text{Tr}\sigma)^2 - 4F(\rho, \sigma)^2}. \quad (234)$$

QED. □

**Lemma 31.** Let  $\Psi^A \in \mathcal{D}(\mathcal{H}_A)$ ,  $\xi^A \in \text{Pos}(\mathcal{H}_A)$  such that  $\|\xi^A - \Psi^A\|_1 \leq \varepsilon$ . Let  $\xi^{AB}, \Psi^{AC}$ ,  $|B| \leq |C|$ , be purifications of  $\xi^A$  and  $\Psi^A$  respectively. Then there exists a partial isometry  $V^{B \rightarrow C}$  such that

$$\|V^{B \rightarrow C} \cdot \xi^{AB} - \Psi^{AC}\|_1 \leq \sqrt{\varepsilon(2 + \varepsilon + 2\sqrt{1 + \varepsilon})}. \quad (235)$$

Note that if it is known that  $\xi^A \in \mathcal{D}_{\leq}(\mathcal{H}_A)$ , then from Corollary 2.2 in Ref. [36], the bound in the RHS can be refined to  $2\sqrt{\varepsilon}$ .

*Proof.* We use the first inequality in the claim of Lemma 30 to have

$$\text{Tr}\xi^A + \text{Tr}\Psi^A - 2F(\xi^A, \Psi^A) \leq \varepsilon. \quad (236)$$

Using the Uhlmann's theorem [37], we claim that there exists a partial isometry  $V^{B \rightarrow C}$  such that  $F(\xi^A, \Psi^A) = F(V^{B \rightarrow C} \cdot \xi^{AB}, \Psi^{AC})$ , and hence,

$$\text{Tr}\xi^{AC} + \text{Tr}\Psi^{AC} - 2F(V^{B \rightarrow C} \cdot \xi^{AB}, \Psi^{AC}) \leq \varepsilon. \quad (237)$$

Since,  $|\text{Tr}\xi^A - \text{Tr}\Psi^A| \leq \varepsilon$ , or,  $\text{Tr}\xi^A \leq 1 + \varepsilon$ , and, using monotonicity,  $F(V^{B \rightarrow C} \cdot \xi^{AB}, \Psi^{AC}) \leq \sqrt{(\text{Tr}\xi^A)(\text{Tr}\Psi^A)} \leq \sqrt{1 + \varepsilon}$ , and hence,  $\text{Tr}\xi^{AC} + \text{Tr}\Psi^{AC} + 2F(V^{B \rightarrow C} \cdot \xi^{AB}, \Psi^{AC}) \leq 2 + \varepsilon + 2\sqrt{1 + \varepsilon}$ . Using the second inequality in the claim of Lemma 30 again, we arrive at

$$\begin{aligned} & \|V^{B \rightarrow C} \cdot \xi^{AB} - \Psi^{AC}\|_1 \\ & \leq \sqrt{[\text{Tr}\xi^{AC} + \text{Tr}\Psi^{AC} - 2F(V^{B \rightarrow C} \cdot \xi^{AB}, \Psi^{AC})][\text{Tr}\xi^{AC} + \text{Tr}\Psi^{AC} + 2F(V^{B \rightarrow C} \cdot \xi^{AB}, \Psi^{AC})]} \\ & \leq \sqrt{\varepsilon(2 + \varepsilon + 2\sqrt{1 + \varepsilon})}. \end{aligned} \quad (238)$$

QED. □

**Corollary 32** (A straightforward corollary of Lemma 9.2 in Ref. [3]). Consider a cq state

$$\rho^{XR} \equiv \sum_{x \in \mathcal{X}} p_x |x\rangle \langle x|^X \otimes \rho_x^R, \quad (239)$$

where  $\rho_x^R \in \mathcal{D}(\mathcal{H}_R)$ ,  $x \in \mathcal{X}$ , and  $\{p_x, x \in \mathcal{X}\}$  is a probability vector. Let  $\rho^R = \text{Tr}_X \rho^{XR}$ ,  $\zeta > 0$ ,  $M \in \mathbb{N}$ , any  $\kappa^R \in \mathcal{D}(\mathcal{H}_R)$ , and  $X^M \equiv (X_1, \dots, X_M)$  be  $M$  i.i.d. random variables with probability distribution  $\{p_x, x \in \mathcal{X}\}$ . Then we have for any  $\alpha \in (1, 2]$ ,

$$\mathbb{E}_{X^M} \left\| \frac{1}{M} \sum_{i=1}^M \rho_{X_i}^R - \rho^R \right\|_1 \leq 4 \exp \left\{ \frac{\alpha - 1}{2\alpha} [\log \nu_{\kappa^R} + D_\alpha(\rho^{XR} \| \rho^X \otimes \kappa^R) - \log M] \right\}. \quad (240)$$

*Proof.* It follows from the claims of Lemma 9.2 in Ref. [3] that for any  $\zeta > 0$ ,

$$\mathbb{E}_{X^M} \left\| \frac{1}{M} \sum_{i=1}^M \rho_{X_i}^R - \rho^R \right\|_1 \leq 2 \sum_x p_x \zeta^{\frac{1-\alpha}{2}} \sqrt{Q_\alpha(\rho_x^R \| \kappa^R)} + \sqrt{\frac{\nu_{\kappa^R} \zeta}{M}} \quad (241)$$

$$= 2\zeta^{\frac{1-\alpha}{2}} \exp \left\{ \frac{\alpha - 1}{2} D_\alpha(\rho^{XR} \| \rho^X \otimes \kappa^R) \right\} + \sqrt{\frac{\nu_{\kappa^R} \zeta}{M}}. \quad (242)$$

If we make a choice of

$$\zeta = \left( \frac{2 \exp \left\{ \frac{\alpha-1}{2} D_\alpha(\rho^{XR} \| \rho^X \otimes \kappa^R) \right\} M}{\nu_{\kappa^R}} \right)^{\frac{2}{\alpha}}, \quad (243)$$

we get

$$\mathbb{E}_{X^M} \left\| \frac{1}{M} \sum_{i=1}^M \rho_{X_i}^R - \rho^R \right\|_1 \leq 4 \exp \left\{ \frac{\alpha-1}{2\alpha} [\log \nu_{\kappa^R} + D_\alpha(\rho^{XR} \| \rho^X \otimes \kappa^R) - \log M] \right\}. \quad (244)$$

QED. □

## C A more general decoupling theorem that we never use!

**Theorem 33.** Let  $\mathcal{X}$  be a finite set,  $\{p_x, x \in \mathcal{X}\}$  a probability distribution on  $\mathcal{X}$ ,  $\rho_x^{AR} \in \mathcal{D}(\mathcal{H}_{AR})$   $\forall x \in \mathcal{X}$ , and  $\{|x\rangle \langle x|^X\}$  a set of orthonormal states in  $X$ . Consider a cq state

$$\rho^{XAR} \equiv \sum_{x \in \mathcal{X}} p_x |x\rangle \langle x|^X \otimes \rho_x^{AR}. \quad (245)$$

For  $M \in \mathbb{N}$ , let  $X_1, \dots, X_M$  be  $M$  independent and identically distributed (i.i.d.) random variables having probability distribution  $\{p_x, x \in \mathcal{X}\}$ , and  $\mathcal{T}^{A \rightarrow E}$  be a class-1 map. Then for  $\alpha \in (1, 2]$ ,  $X_1^M \equiv (X_1, \dots, X_M)$ , random Unitaries  $U_1^M \equiv (U_1, \dots, U_M)$  acting independently on  $A$ , we have for any  $\sigma^R, \kappa^R \in \mathcal{D}(\mathcal{H}_R)$ ,

$$\begin{aligned} \mathbb{E}_{X_1^M} \mathbb{E}_{U_1^M} \left\| \frac{1}{M} \sum_{i=1}^M \mathcal{T}(U_i \cdot \rho_{X_i}^{AR}) - \omega_{\mathcal{T}}^E \otimes \rho^R \right\|_1 \\ \leq 4 \exp \left\{ \frac{\alpha-1}{2\alpha} [\log \nu_{\sigma^R} + D_\alpha(\rho^{XAR} \| \rho^X \otimes \mathbb{1}^A \otimes \sigma^R) - \log M + \Theta(\mathcal{T})] \right\} \text{ind}_{|A| \neq 1} \\ + 4 \exp \left\{ \frac{\alpha-1}{2\alpha} [\log \nu_{\kappa^R} + D_\alpha(\rho^{XR} \| \rho^X \otimes \kappa^R) - \log M] \right\} \text{ind}_{|\mathcal{X}| \neq 1}. \end{aligned} \quad (246)$$

*Proof.* We have

$$\begin{aligned} \mathbb{E}_{X_1^M} \mathbb{E}_{U_1^M} \left\| \frac{1}{M} \sum_{i=1}^M \mathcal{T}(U_i \cdot \rho_{X_i}^{AR}) - \omega_{\mathcal{T}}^E \otimes \rho^R \right\|_1 \\ \leq \mathbb{E}_{X_1^M} \mathbb{E}_{U_1^M} \left\| \frac{1}{M} \sum_{i=1}^M [\mathcal{T}(U_i \cdot \rho_{X_i}^{AR}) - \omega_{\mathcal{T}}^E \otimes \rho_{X_i}^R] \right\|_1 + \mathbb{E}_{X_1^M} \left\| \frac{1}{M} \sum_{i=1}^M \omega_{\mathcal{T}}^E \otimes \rho_{X_i}^R - \omega_{\mathcal{T}}^E \otimes \rho^R \right\|_1 \\ = \mathbb{E}_{X_1^M} \mathbb{E}_{U_1^M} \left\| \frac{1}{M} \sum_{i=1}^M [\mathcal{T}(U_i \cdot \rho_{X_i}^{AR}) - \omega_{\mathcal{T}}^E \otimes \rho_{X_i}^R] \right\|_1 \text{ind}_{|A| \neq 1} + \\ \mathbb{E}_{X_1^M} \left\| \frac{1}{M} \sum_{i=1}^M \rho_{X_i}^R - \rho^R \right\|_1 \text{ind}_{|\mathcal{X}| \neq 1}, \end{aligned} \quad (247)$$

where the inequality follows from the triangle inequality and the last equality follows since  $\|X \otimes Y\|_1 = \|X\|_1 \|Y\|_1$ , and the first and the second terms are identically zero if  $|A| = 1$  and  $|\mathcal{X}| = 1$  respectively. The upper bound for the second term can be deduced from Lemma 9.2 in Ref. [3] for any  $\alpha \in (1, 2]$  and any  $\kappa^R \in \mathcal{D}(\mathcal{H}_R)$  as

$$\mathbb{E}_{X_1^M} \left\| \frac{1}{M} \sum_{i=1}^M \rho_{X_i}^R - \rho^R \right\|_1 \leq 4 \exp \left\{ \frac{\alpha - 1}{2\alpha} [\log \nu_R + D_\alpha(\rho^{XR} \| \rho^X \otimes \kappa^R) - \log M] \right\}. \quad (248)$$

Note that Lemma 9.2 in Ref. [3] doesn't provide an upper bound in the above form but it is easy to deduce it from the claim, and, for the sake of completeness, it is provided in Corollary 32.

The rest of the proof is to upper bound the first term in (247). For  $\zeta > 0$  and  $\forall x \in \mathcal{X}$ , let  $\Pi_x^{AR} \equiv \{\mathcal{M}_{\mathbb{1}^A \otimes \sigma^R}(\rho_x^{AR}) \geq \zeta \mathbb{1}^A \otimes \sigma^R\}$ ,  $\hat{\Pi}_x^{AR} \equiv \mathbb{1}^{AR} - \Pi_x^{AR}$ ,  $\mu_{1,x} \equiv \omega_{\mathcal{T}}^E \otimes \text{Tr}_A \{\Pi_x^{AR} \rho_x^{AR}\}$ , and  $\mu_{2,x} \equiv \omega_{\mathcal{T}}^E \otimes \text{Tr}_A \{\hat{\Pi}_x^{AR} \rho_x^{AR}\}$ . Note that  $\mu_{1,x} + \mu_{2,x} = \omega_{\mathcal{T}}^E \otimes \rho_x^R$ . We now have from the triangle inequality

$$\begin{aligned} \mathbb{E}_{X_1^M} \mathbb{E}_{U_1^M} \left\| \frac{1}{M} \sum_{i=1}^M [\mathcal{T}(U_i \cdot \rho_{X_i}^{AR}) - \omega_{\mathcal{T}}^E \otimes \rho_{X_i}^R] \right\|_1 \\ \leq \mathbb{E}_{X_1^M} \mathbb{E}_{U_1^M} \left\| \frac{1}{M} \sum_{i=1}^M \{\mathcal{T}[U_i \cdot (\Pi_{X_i}^{AR} \rho_{X_i}^{AR})] - \mu_{1,X_i}\} \right\|_1 + \\ \mathbb{E}_{X_1^M} \mathbb{E}_{U_1^M} \left\| \frac{1}{M} \sum_{i=1}^M \{\mathcal{T}[U_i \cdot (\hat{\Pi}_{X_i}^{AR} \rho_{X_i}^{AR})] - \mu_{2,X_i}\} \right\|_1. \end{aligned} \quad (249)$$

We attack the first term.

$$\begin{aligned} \mathbb{E}_{X_1^M} \mathbb{E}_{U_1^M} \left\| \frac{1}{M} \sum_{i=1}^M \{\mathcal{T}[U_i \cdot (\Pi_{X_i}^{AR} \rho_{X_i}^{AR})] - \mu_{1,X_i}\} \right\|_1 \\ \leq \mathbb{E}_{X_1^M} \mathbb{E}_{U_1^M} \left\| \frac{1}{M} \sum_{i=1}^M \mathcal{T}[U_i \cdot (\Pi_{X_i}^{AR} \rho_{X_i}^{AR})] \right\|_1 + \mathbb{E}_{X_1^M} \|\mu_{1,X_i}\|_1 \end{aligned} \quad (250)$$

$$\leq \frac{2}{M} \sum_{i=1}^M \mathbb{E}_{X_i} \mathbb{E}_{U_i} \left\| \mathcal{T}[U_i \cdot (\Pi_{X_i}^{AR} \rho_{X_i}^{AR})] \right\|_1 = 2 \mathbb{E}_X \mathbb{E}_U \left\| \mathcal{T}[U \cdot (\Pi_X^{AR} \rho_X^{AR})] \right\|_1 \quad (251)$$

$$\leq 2 \mathbb{E}_X \left\| \Pi_X^{AR} \rho_X^{AR} \right\|_1 = 2 \sum_x p_x \left\| \Pi_x^{AR} \rho_x^{AR} \right\|_1 \quad (252)$$

$$\leq 2\zeta^{\frac{1-\alpha}{2}} \sum_x p_x \sqrt{Q_\alpha(\rho_x^{AR} \| \mathbb{1}^A \otimes \sigma^R)} \quad (253)$$

$$= 2\zeta^{\frac{1-\alpha}{2}} \exp \left\{ \frac{\alpha - 1}{2} D_\alpha(\rho^{XAR} \| \rho^X \otimes \mathbb{1}^A \otimes \sigma^R) \right\}, \quad (254)$$

where the first inequality follows from the triangle inequality, the second inequality follows from the convexity of the trace norm to have

$$\begin{aligned} \mathbb{E}_{X_1^M} \|\mu_{1,X_i}\|_1 &= \mathbb{E}_{X_1^M} \left\| \frac{1}{M} \sum_{i=1}^M \mathbb{E}_{U_i} \left\{ \mathcal{T} [U_i \cdot (\Pi_{X_i}^{AR} \rho_{X_i}^{AR})] \right\} \right\|_1 \\ &\leq \frac{1}{M} \sum_{i=1}^M \mathbb{E}_{X_i} \left\| \mathbb{E}_{U_i} \left\{ \mathcal{T} [U_i \cdot (\Pi_{X_i}^{AR} \rho_{X_i}^{AR})] \right\} \right\|_1 \leq \frac{1}{M} \sum_{i=1}^M \mathbb{E}_{X_i} \mathbb{E}_{U_i} \left\| \mathcal{T} [U_i \cdot (\Pi_{X_i}^{AR} \rho_{X_i}^{AR})] \right\|_1, \end{aligned} \quad (255)$$

and similarly for the first term, the first equality follows since  $X_i$ 's and  $U_i$ 's are i.i.d., the third inequality follows from the definition of class-1 maps, the fourth inequality follows from Lemma 28 (proved by Hayashi [3]), and the last inequality follows from the concavity of  $x \mapsto \sqrt{x}$ .

We now attack the second term. Let  $\Delta_{X_i U_i} \equiv \mathcal{T} [U_i \cdot (\hat{\Pi}_{X_i}^{AR} \rho_{X_i}^{AR})] - \mu_{2,X_i}$  and  $\Delta_{X_1^M U_1^M} \equiv \sum_{i=1}^M \Delta_{X_i U_i} / M$ . Note that  $\mathbb{E}_{X_1^M U_1^M} \left\{ \Delta_{X_i U_i} \Delta_{X_j U_j}^\dagger \right\} = \mathbf{0}, \forall i \neq j$ , and hence,

$$\mathbb{E}_{X_1^M U_1^M} \left\{ \Delta_{X_1^M U_1^M} \Delta_{X_1^M U_1^M}^\dagger \right\} = \frac{1}{M^2} \sum_{i=1}^M \mathbb{E}_{X_i U_i} \left\{ \Delta_{X_i U_i} \Delta_{X_i U_i}^\dagger \right\} = \frac{1}{M} \mathbb{E}_{XU} \left\{ \Delta_{XU} \Delta_{XU}^\dagger \right\} \quad (256)$$

$$\leq \frac{|A|^2 \text{Tr}_{A'} (\omega_{\mathcal{T}}^{E_{A'}})^2}{M(|A|^2 - 1)} \otimes \text{Tr}_A \mathbb{E}_X \left\{ \hat{\Pi}_X^{AR} (\rho_X^{AR})^2 \hat{\Pi}_X^{AR} \right\}, \quad (257)$$

where the inequality follows from Lemma 27. Following the arguments in Theorem 1 in dealing with the second term, we get

$$\mathbb{E}_{X_1^M U_1^M} \left\| \frac{1}{M} \sum_{i=1}^M \left\{ \mathcal{T} [U_i \cdot (\hat{\Pi}_{X_i}^{AR} \rho_{X_i}^{AR})] - \mu_{2,X_i} \right\} \right\|_1 \leq \sqrt{\frac{\nu_{\sigma^R} \zeta |A|^2 \exp \{ \Theta(\mathcal{T}) \}}{M(|A|^2 - 1)}}. \quad (258)$$

We now have

$$\begin{aligned} \mathbb{E}_{X_1^M U_1^M} \left\| \frac{1}{M} \sum_{i=1}^M [\mathcal{T}(U_i \cdot \rho_{X_i}^{AR}) - \omega_{\mathcal{T}}^E \otimes \rho_{X_i}^R] \right\|_1 \\ \leq 2\zeta^{\frac{1-\alpha}{2}} \exp \left\{ \frac{\alpha-1}{2} [D_\alpha(\rho^{XAR} \| \rho^X \otimes \mathbb{1}^A \otimes \sigma^R)] \right\} + \sqrt{\frac{\nu_{\sigma^R} \zeta |A|^2 \exp \{ \Theta(\mathcal{T}) \}}{M(|A|^2 - 1)}}, \end{aligned} \quad (259)$$

and by appropriately choosing  $\zeta$ , we get

$$\begin{aligned} \mathbb{E}_{X_1^M U_1^M} \left\| \frac{1}{M} \sum_{i=1}^M [\mathcal{T}(U_i \cdot \rho_{X_i}^{AR}) - \omega_{\mathcal{T}}^E \otimes \rho_{X_i}^R] \right\|_1 \\ \leq 4 \exp \left\{ \frac{\alpha-1}{2\alpha} [\log \nu_{\sigma^R} + D_\alpha(\rho^{XAR} \| \rho^X \otimes \mathbb{1}^A \otimes \sigma^R) - \log M + \Theta(\mathcal{T})] \right\}. \end{aligned} \quad (260)$$

The claim now follows from (247), (248), and (260).  $\square$



## References

- [1] T. M. Cover and J. A. Thomas. *Elements of information theory*. Wiley, Hoboken, NJ, USA, 2nd edn., 2006.
- [2] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, Cambridge, 2000.
- [3] M. Hayashi. *Quantum information: An introduction*. Springer, Berlin, 2006.
- [4] M. Ohya and D. Petz. *Quantum Entropy and its use*. Springer-Verlag, Berlin, 1st edn., 1993.
- [5] M. M. Wilde. *Quantum information theory*. Cambridge University Press, 2013.
- [6] R. Renner. *Security of quantum key distribution*. Ph.D. thesis, Swiss Federal Institute of Technology, ETH, Zurich, Switzerland, 2005.
- [7] M. Tomamichel. *A framework for non-asymptotic quantum information theory*. Ph.D. thesis, Swiss Federal Institute of Technology, ETH, Zurich, Switzerland, 2012.
- [8] R. G. Gallager. *A simple derivation of the coding theorem and some applications*. *IEEE Trans. Inf. Theory*, vol. 11: pp. 3–18, Jan. 1965.
- [9] R. G. Gallager. *Information theory and reliable communication*. John Wiley & Sons, Inc., New York, 1968.
- [10] V. Strassen. *Asymptotische abschätzungen in Shannon’s informationtheorie*. In *Trans. third Prague Conf. Inf. Theory*, pp. 689–723. 1962.
- [11] M. V. Burnashev and A. S. Holevo. *On reliability function of quantum communication channel*. *Prob. Inf. Trans.*, vol. 34: pp. 97–107, 1998.
- [12] A. S. Holevo. *Reliability function of general classical-quantum channel*. *IEEE Trans. Inf. Theory*, vol. 46: pp. 2256–2261, 2000.
- [13] M. Hayashi. *Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding*. *Phys. Rev. A*, vol. 76: p. 062301, 2006.
- [14] I. Devetak, A. W. Harrow, and A. Winter. *A family of quantum protocols*. *Phys. Rev. Lett.*, vol. 93: p. 230504, Dec 2004.
- [15] I. Devetak, A. W. Harrow, and A. Winter. *A resource framework for quantum Shannon theory*. *IEEE Trans. Inf. Theory*, vol. 54: pp. 4587–4618, 2008.
- [16] B. Schumacher and M. A. Nielsen. *Quantum data processing and error correction*. *Phys. Rev. A*, vol. 54: pp. 2629–2635, Oct. 1996.

- [17] F. Dupuis. *The decoupling approach to quantum information theory*. Ph.D. thesis, University of Montreal, Montreal, Canada, 2009.
- [18] P. Hayden. *Decoupling: A building block for quantum information theory*. Tutorial at Quantum Information Processing (QIP) workshop, Jan. 2011.
- [19] P. Hayden, M. Horodecki, A. Winter, and J. Yard. *A decoupling approach to the quantum capacity*. *Open Syst. Inf. Dyn.*, vol. 15: pp. 7–19, 2008.
- [20] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter. *The mother of all protocols: restructuring quantum information's family tree*. *Proc. R. Soc. A*, vol. 465: pp. 2537–2563, 2009.
- [21] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner. *One-shot decoupling*. *Commun. Math. Phys.*, vol. 328: pp. 251–284, 2014.
- [22] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel. *On quantum Rényi entropies: a new generalization and some properties*. *J. Math. Phys.*, vol. 54: p. 122203, 2013.
- [23] M. M. Wilde, A. Winter, and D. Yang. *Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy*. *Commun. Math. Phys.*, vol. 331: pp. 593–622, Oct. 2014.
- [24] D. Petz. *Quasi-entropies for finite quantum systems*. *Rep. Math. Phys.*, vol. 23: pp. 57–65, Feb. 1986.
- [25] D. Petz. *From  $f$ -divergence to quantum quasi-entropies and their use*. *Entropy*, vol. 12: pp. 304–325, Mar. 2010.
- [26] M. Tomamichel, M. Berta, and M. Hayashi. *Relating different quantum generalizations of the conditional Rényi entropy*. *J. Math. Phys.*, vol. 55: p. 082206, 2014.
- [27] B. Schumacher. *Quantum coding*. *Phys. Rev. A*, vol. 51: pp. 2738–2747, 1995.
- [28] I. Csiszár and J. Körner. *Information Theory: Coding theorems for discrete memoryless systems*. Cambridge University Press, New York, USA, 2nd edn., 2011.
- [29] I. Devetak. *Triangle of dualities between quantum communication protocols*. *Phys. Rev. Lett.*, vol. 97: p. 140503, 2006.
- [30] M. Horodecki, J. Oppenheim, and A. Winter. *Quantum state merging and negative information*. *Commun. Math. Phys.*, vol. 269: pp. 107–136, 2007.
- [31] C. A. Fuchs and J. van de Graaf. *Cryptographic distinguishability measures for quantum-mechanical states*. *IEEE Trans. Inf. Theory*, vol. 45: pp. 1216–1227, May 1998.

- [32] I. Devetak and J. Yard. *Exact cost of redistributing multipartite quantum states*. *Phys. Rev. Lett.*, vol. 100: p. 230501, Jun 2008.
- [33] M.-Y. Ye, Y.-K. Bai, and Z. D. Wang. *Quantum state redistribution based on a generalized decoupling*. *Phys. Rev. A*, vol. 78: p. 030302, 2008.
- [34] B. Groisman, S. Popescu, and A. Winter. *Quantum, classical, and total amount of correlations in a quantum state*. *Phys. Rev. A*, vol. 72: p. 032317, 2005.
- [35] R. Bhatia. *Matrix Analysis*. Springer-Verlag, New York, 1997.
- [36] F. Dupuis, O. Szehr, and M. Tomamichel. *A decoupling approach to classical data transmission over quantum channels*. *IEEE Trans. Inf. Theory*, vol. 60: pp. 1562–1572, Mar. 2014.
- [37] A. Uhlmann. *The ‘transition probability’ in the state space of a  $*$ -algebra*. *Rep. Math. Phys.*, vol. 9: pp. 273–279, 1976.